Reviewer’s report

Title: The short international physical activity questionnaire: cross-cultural adaptation, validation and reliability of the Hausa language version in Nigeria

Version: 1 Date: 22 February 2011

Reviewer: Michael Erhart

Reviewer’s report:

Review of the manuscript „The short international physical activity questionnaire: cross-cultural adaptation, validation and reliability of the Hausa language version in Nigeria.

The above manuscript describes the process of the cultural and language adaptation of the English IPAQ short physical activity questionnaire into Hausa language in Nigeria. Psychometric testing results issued from 102 respondents were reported as well. With increasing trend in international epidemiological research activities there is an increasing need for cross-cultural applicable instruments. The current study thus has public health relevance. The adaptation could add to strengthen the armamentarium of available health questionnaires. However there are some aspects connected with the actual manuscript that had to be viewed critically. Generally spoken the process of cultural adaptation was performed very well and with adequate methods applied in a straightforward manner. However the testing of the psychometric analyses are less satisfying:

Major Compulsory Revisions

General comments:

1. The actual sample is rather small and the respondents do not seem to show sizeable variation in their health status. Hopefully there was enough variation in their physical activity (the age range from 20-65 hints at this). Unfortunately only few information on the respondents health and status and physical fitness was gathered – not even information on body-mass index was mentioned. There

2. The description of the methods is rather sparse – e.g. it remains unclear how the instrument is actually scored: How are the test scores issued from the item responses. No information is provided on the actual metric of the test score theoretically (minimal and maximal possible value) There are results mentioned on the total score, however no information is presented how such a total score is calculated from the individual items. Are the minutes of very hard work simply added with the minutes of hard work? Are very hard work minutes upweighted?

3. If however a total score is calculated, then it would be interesting to test how well the different items could be combined into an overall score (e.g. item-total correlation (corrected or overlap as is the default with spss 15), cronbach alpha coefficient for internal consistency of item responses could be also calculated).
Of course I do understand that in the case of the IPAQ the classical psychometric test-model might not be adequate for the instrument: A high level of very hard physical activities might not necessarily be associated with high level of walking. Thus one would not necessarily assume that the item-scores should display a consistent picture. Based on such considerations a rationale for the actual statistical analyses could be given.

4. Connected with point 2 and 3 I think that it is necessary to give some explanations with regards to the theoretical assumptions underlying the theoretical measurement model of the IPAQ. For example how does the information on sitting time add to the measurement of physical activity?

5. I wonder some more about the question on sitting time: What about a person who performs hard physical activities in a sitting position? It is not necessary to stress the picture of e.g. a body-builder who lifts e.g. 100-150 pounds several times first in lying position and then in a sitting position to understand that the items are not that exclusive that the minutes could be simply added.

An extreme example is a machine used for harvesting some specific vegetables in specific regions in Europe: people are lying in row on the wings of a device that is slowly moved. Lying head down these person pick up the vegetables and put them on the machine (a hard work in a lying position!)

6. The choice of statistical analyses is not consistent on the one hand spearman correlation are calculated (nonparametric statistic appropriate for ordinal and non-normal variables) on the other hand ICCs are calculated that implicitly assumes interval-scaled variables and normal distributed values. (as would item-total correlation and Cronbach alpha coefficient). From a theoretical point of view the items assess information on an interval scaled level (there even is a natural zero). Given the test scores are approximately normal distributed, parametric statistics could be calculated.

7. The actual psychometric analyses are rather sparse: no information is provided on the empirical minimal and maximal observed scores, score frequency distribution, shape of the test score distribution. These aspects could have been examined and compared across test-versions as well. The authors did also not take into account or tested for gender differences, differences in socioeconomic status or other group differences. Especially different work environments (the authors do have this information in their data) could be expected to result in different levels of physical activity.

Specific comments:

Page 3, line 1-3: It is not possible to conclude that the reliability and concurrent validity of Hausa IPAQ is comparable to that of the original English IPAQ: This statement is only valid if it had been explicitly tested that e.g. the retest reliability of the Hausa version is similar to the retest reliability of the English version. Concurrent validity is of course similar for both. Better try to interpret the magnitude of the concurrent validity coefficients.
Page 7, para 1 and page 9, para 2: Please provide some more information on this cognitive debriefing of the instrument: Were the interviews performed in a structured way e.g., the interviewer asked a set of prespecified questions or in a narrative way? It even remain unclear if the respondents were questioned and interviewed one by one or if a group discussion was enabled?

Page 11, para 2: I wonder if the item on sitting time is understood in the same way as it is by e.g., North American respondents. Hard physical activities could be also performed in a sitting position (in North America also of course but maybe less frequent).

Page 12, line 1-4: It remains absolutely unclear if construct validity of the IPAQ was assessed with regards to these aspects in the current study as well? If yes – these results should be reported. If not – please provide a rationale why this wasn’t assessed in the current study as well.

Page 14-15 Conclusion: In its actual form the conclusion is a mere summary of the results. Please try to rewrite and focus on a more comprehensive evaluation of the measure: For which purpose under which circumstances the adapted IPAQ could be employed. And if employed what had to bear in mind – what are the shortcomings of the measure? What further development or testing activities are warranted? E.g., assessing the psychometric properties of the measure in a large sample; further assessment of validity using valid information on actual physical activity and fitness; assessing the sensitivity to change; etc.

Minor Essential Revisions

Page 2, abstract, line 1-4: Awkwardly formulated sentence. Try to make to sentences out of it.

Discretionary Revisions

none

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests