Reviewer’s report

Title: Computer Adaptive Assessment of Individual Mental Perceptions of Job Contents on a Website for Workplace Employees

Version: 1 Date: 10 November 2010

Reviewer: Giulio Vidotto

Reviewer’s report:

The paper presents an interesting application of an adaptive tool to assess job satisfaction. The possibility to use a skewness analysis approach instead of classical procedures to analyze differences among groups is explored. Despite the interesting contents and the strengths of the study, several improvements are needed in order to consider the paper for publication.

The authors provide an alternative way to analyze differences in items among subjects across years. This alternative methodology seems to need a formal, methodological demonstration and explanation. In other words, if this methodology - as it seems - is here introduced for the first time, some comparative analyses to evaluate its goodness (with a formal analytical proof of its coherence) are required. On the other hand, if this formal approach has already been explored by previous literature, the references have to be considered and discussed. On the bottom of page 10 authors assert: “Previous studies found that skewness analysis is useful in evaluating dissimilarity of examinee groups”, but no reference is here reported.

The way authors describe their study prevents us from fully understanding both the steps of the procedure as well as the discussion. This is particularly due to a very puzzling exposition of the results. Furthermore, many methodological issues are unclear and potentially inappropriate.

The impression obtained out of the paper is a miscellaneous of procedures and analyses only partially supported by a well-organized project and strong references in the literature.

The opening question is quite well exposed by the authors, nevertheless the description of the applied methods as well as the exposition of the results are extremely cryptic. For all these reasons it is difficult to understand the interpretation of both the results and the related decisions suggested by the authors.

The paper adheres to the relevant standards for reporting and data deposition. Nevertheless, the introduction of a paragraph to describe the use of the Confidence Intervals with the Cronbach alpha seems redundant.

Finally, abstract has to be upgraded according to the improvements to be introduced in the methods, results and discussion paragraphs.

- Major Revisions
1) The “Standardized assessments of health status” paragraph has to be improved since it seems an oversimplification of the state of art in the field. Authors should clearly cite the SCID instead of speaking of a time consuming structured interview. The review of the state of art has to be improved and extended. Furthermore, it has to be justified how the short form of questionnaires have raised concern, since there are many cases in the literature where the reduced form of questionnaires present excellent levels of reliability and validity.

2) Methods are presented in a quite puzzling way. It is not clear which group performed the CAT assessment and which one the traditional pen and pencil form. I can guess that 2008 group has been used to calibrate the parameters of the procedure and 2009 data have been later collected through the adaptive tool. Anyway, if so, it has to be better explained. Furthermore, several aspects of the comparison between 2008 and 2009 have to be reformulated considering the different methodologies for data collection.

3) Chronbach’s alpha coefficients provided for the whole instrument in 2008 and 2009 are extremely high. In these cases a very high value could be a critical rather than a strong point of the tool. Further investigation and interpretation are recommended. Moreover, it is not clear at all the kind of procedure adopted for the point-biserial correlation analysis. In the literature this kind of correlation is calculated between a dichotomous and a continuous variable. It is necessary to know how authors dichotomized the supposed 4-point scale, or - if a different procedure was used - it is essential to know it.

4) No descriptive statistics concerning the adaptive assessment are provided. For instance, how many items (in average) were administered to participants in order to reach the stop criterion? Additionally, no information is provided about the algorithm used to select the item on the basis of the previous responses of the participant. This information is crucial in order to evaluate the reliability and the efficiency of the assessment.

5) The construction of the 37 items version of the questionnaire is unclear. No information is provided for the 15 items added to the 22 items of the Chinese version of JCQ.

6) Figures 1, 2 and 3 have to be explained. In present format, it is really difficult to understand them. More specifically and given what emerges in the figures, conclusions resulting from figure 1 are at least arbitrary. Considering figure 2, the selection of the groups investigated is made without any explanation (i.e. why the category “>56 years” is not included? This is in contrast with figure 3). No reference point is present in the figures and this aspect makes extremely unclear and complicated the interpretation. The basis of the considerations carried out on groups 9 and 18 in figure 2 have to be disclosed.

7) The paragraph “Web-CAT performance” has to be deeply modified in order to make it understandable and evaluable. More specifically, neither the step 2 nor the step 3 of the procedure is clear. It is not specified what considerations have been done concerning “aberrant responses” (were they discarded?).

8) The conclusion that -0.37 could overcome the shortcomings of traditional assessments has to be further discussed.
9) Concerning the skewness analysis, authors have to clarify the advantages of this method. If any significant improvement is present it should be better argued, explained and formalized. This crucial aspect of the study should be better explored.

10) It seems reasonable to introduce a more accurate explanation about CAT basis and computations. Indeed, these steps are crucial for a reader to understand and evaluate the whole procedure, the results, and - above all - the conclusions.

**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