Author's response to reviews

Title: Web-based computer adaptive assessment of individual perceptions of job satisfaction for hospital workplace employees

Authors:

Tsair-Wei Chien (smile@mail.chimei.org.tw)
Wen-Pin Lai (shumiin@seed.net.tw)
Chih-Wei Lu (chwelu@gmail.com)
Shih-Chung Chen (chung@mail.stut.edu.tw)
Shih-Bin Su (shihbin.su@msa.hinet.net)

Version: 2 Date: 18 November 2010

Author's response to reviews: see over
Cover letter

Thank you very much for giving us another opportunity to revise our manuscript titled “Web-based computer adaptive assessment of individual perceptions of job satisfaction for hospital workplace employees” for reconsideration for publication in *BMC Medical Research Methodology*. My co-authors and I very much appreciate the thoughtful comments and suggestions made by the reviewers to help us improve our work. Accordingly, we have substantially revised the manuscript and we hope that we have addressed all of the concerns raised by the reviewers. We have summarized our responses point-by-point in the “Response to Reviewers’ Comments” section starting on the next page.

Briefly, we have simplified Table 2 according to the reviewer’s suggestion regarding unidimensionality and reorganized the manuscript structure for ease of understanding. This revision has also been professionally edited, as suggested, to improve its readability.

We remain confident that our study described in this revised manuscript is of great value and interest to the readership of *BMC Medical Research Methodology*. Your reconsideration for its publication would be very much appreciated.

Regarding ethics, we have highlighted the statement in the methods section that granted ethical approval for this study.

Author’s response to reviews

Reviewer's report
Title: Computer Adaptive Assessment of Individual Mental Perceptions of Job Contents on a Website for Workplace Employees
Version: 1 Date: 15 June 2010
Reviewer: Cynthia Leung

Major compulsory revisions:
1. The authors described two research questions. The first one was on the feasibility of applying skewness coefficient analyses, instead of traditional t test, ANOVA or correlation to compare perception changes between groups across items over years. While the authors have described in details the skewness coefficient results, it will also be useful to compare the results with the t test, ANOVA and correlation results, and to discuss the similarities and discrepancies,
and explain the discrepancies. This will enable the reader to appreciate the advantages of using skewness coefficients in analysis. While it is understandable to identify worker concerns, the rationale or significance of perception change over time should also be explained.

**Response:**

As suggested by the reviewer, we have addressed the advantage of skewness usage not only for the graphical representation but also for the reason of truth that data points are not perfectly symmetric in reality. It might be problematic to compare results with the t test and ANOVA. Due to space limit, we only recommended readers to referring preceding published paper [24] where some discussions and descriptions were reported.

In addition, we identified worker concerns and main findings in the revised discussion session to stress that (1) workers aged 26 - 35 felt that job content satisfaction was becoming significantly worse than the previous year; and (2) two groups, age 18-25 and high school education, showed a significant deterioration on item 8 from 2008 to 2009.

2. The second research question was whether a CAT based self-evaluation questionnaire can facilitate a more useful and faster response from institutional services. This manuscript has not addressed this research question.

**Response:**

Because many researches have verified the advantage of CAT against NAT(non-adaptive testing), we have not particularly addressed the issue in the manuscript, instead focusing on the application of polytomous CAT assessment on Internet.

3. The authors used infit and outfit statistics to examine unidimensionality. Another commonly used procedure to assess dimensionality is principal component analysis of the residuals. These results should also be reported.

**Response:**

In response to the reviewer’s comment, we also referred to Tennant and Pallant study[34] and Richard Smith's suggestions[35] that exploratory factor analysis (EFA), especially using parallel analysis (PA)[36], should be undertaken to make sure dimensionality of study data. These results with regard to checking dimensionality have reported in the revised manuscript.

4. The questionnaire is a 4-point likert type scale. Category functioning should be examined to see whether the participants could meaningfully distinguish between
the categories. These results should also be reported.

Response:

As suggested by the reviewer, we have added following sentence in results:

step threshold difficulties under the Rasch rating scale model for the 24-item questionnaire were -4.16, -1.50 and 2.66, which is congruent with the guidelines proposed by Linacre [49] that (1) that average measures advance monotonically within each category; (2) step calibrations advance; (3) that step difficulties advance by at least 1.4 logits; and (4) that step difficulties advance by less than 5.0 logits.

5. The authors described a cut-off point of -0.37 logits but gave the impression that this was based on item 14. If the JCQ-37 is a unidimensional measure and the individual items scores can be added up to form a total score, it would make sense to use the total score to form a cut-off, instead of one item.

Response:

Due to space limit, we have removed the issue of cut-off point discussed in the original manuscript and focused more attention on the two research questions of Web-CAT application and skewness analysis feasibility.

6. Normally, in the calculation of sensitivity and specificity, one compares the test results with some external “gold standard”. In this particular case, it is not clear what the external gold standard is.

Response:

As mentioned as previous comment and response, the setting of cut-off point and the calculation of sensitivity and specificity needs more sophisticated study and exploration for the job satisfaction survey. It is worth carrying out the research in future.

7. Though the authors described differences in job satisfaction among various groups in the results section, these points were not discussed in the discussion section.

Response:

Besides focusing attention on the two research questions of Web-CAT application and skewness analysis feasibility, we highlighted the main findings of this study regarding differences in job satisfaction among various groups in revised discussion session for further advanced concerns. There are (1) workers aged 26 - 35 felt that job content satisfaction was becoming significantly worse than the previous year; and (2)
groups, age 18-25 and high school education, showed a significant deterioration on item 8 (salary and wage levels compared with other hospitals are) from 2008 to 2009.
Reviewer's report:

- Major Compulsory Revisions
  None.

- Minor Essential Revisions
  1) In 3 points of the paper (In METHODS - Steps of instrument selection - Validation procedure - (3) Web-CAT assessment: last line of the paragraph; in RESULTS - Web-CAT performance: last line of the paragraph; in Table 2: last line under the table) you say the JCQ-37 Web-CAT questionnaire is available at http://www.healthup.org.tw/irt_test4/irt_start.htm. When I try to access this webpage, I get what seems to be an error message page (since I can't read Chinese characters I attach a screenshot for you to check, Error-message.jpg): I suggest that you check the above-mentioned URL and, if the link is effectively broken, that you fix the problem (or, if it wouldn't be possible, I suggest that you remove the URLs from the text).

**Response:**

In response to the reviewer’s comment, we have fixed the Web-CAT module on Internet. That is one of the features in this study for readers to practice in their wish. We also added Figure 1 to describe the CAT flowchart and procedure designed in this study.

3) All the references citations follow the punctuation marks in the text (e.g. in INTRODUCTION, line 5: "Optimizing the way in which healthcare providers use institutional services to maximize likelihood of health promotion outcomes is urgent and essential. [6,7]").

I think the paper legibility would improve if these citations came before punctuation marks.

**Response:**

As suggested by the reviewer, we have improved all the references citations that should follow, instead before, the punctuation marks to conform to the journal style.

2) In Figure legends section, please change "Figure 1:" into "Figure 1."

**Response:**

As suggested by the reviewer, we have changed the figure legends to be correct.
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.

Response:

As suggested by the reviewer, we have added the reference [24] to the revised manuscript. Readers who are interested in the skewness analysis in group comparison can practice it with an Excel-VBA module. In the discussion of the previous published paper, there are detailed discussion of the similarities and discrepancies against t-test and ANOVA.

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.

Response:

We have made substantial revision to the manuscript to hopefully address every concern matter of the research questions as well as reorganized the manuscript structure for ease of understanding. This revision has also been professionally edited, as suggested, to improve its readability.

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.

**Response:**

As commended by the reviewer, we have made a lot changes in the revised manuscript, especially on the two research questions of Web-CAT application and skewness analysis feasibility. Through which, readers could be easily understood to the revised version of manuscript.

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.

**Response:**

As suggested by the reviewer, we have removed the topic of Cronbach alpha and its Confidence Intervals and paid more attention on the two main research questions of Web-CAT application and skewness analysis feasibility. As for the abstract, we also revised it to be readable and understandable than previous version of manuscript.

- 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.

**Response:**

In response to the reviewer’s comment, we concerns about not only the reliability and validity of assessments but also the efficiency and detection of aberrant responses from respondents, especially in the computerization era.

With the right item bank and a tailored individual assessment, CAT can be much more efficient than a traditional paper-and-pencil test, since a test which is either very easy or very difficult provides relatively little information about the examinee's ability level. On the other hand, using CAT the examinees can be given the items that
maximize the information about their ability levels from the item responses. Consequently, examinees will be given fewer test items that are very easy or very hard for them to complete. This tailored item selection can result in reduced standard errors and greater precision with merely a handful of properly selected items. That is the reason we proposed in the manuscript. The CAT can be another alternative besides the lengthy and shorten forms of assessment in clinical settings.

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.

Response:
As suggested by the reviewer, we have reorganized and clarified the method description in the revised manuscript. The self-administered 37-item Job Content Questionnaire (JCQ-37) was designed for use on a website with assessments of NAT (non adaptive testing) and CAT in 2008 and 2009, respectively. Two illustrations were raised in Figure 1 and 2 to present the results using person measures and specific item (e.g., item 8) in compassions of groups on perception of job satisfaction for hospital workers.

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.

Response:
In response to the reviewer’s comment, we have removed the session regarding Chronbach’s alpha and point-biserial correlation analysis because those topics would mislead readers’ attention to the two main research questions of Web-CAT application and skewness analysis feasibility.

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.

**Response:**

As suggested by the reviewer, we added Figure 1 to present the flowchart and procedure of CAT implementation. The relevant information regarding CAT design, for instance, stop rule, item selection method and person estimation algorithm has been described in great detail in Methods and Figure 1.

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.

**Response:**

In response to the reviewer’s comment, we have clarified the questionnaire data derived from eight items (related to supervisors and coworkers supports in Chinese Version of the JCQ) and 29 items regarding job satisfaction to form the 37-item Job Content Questionnaire (JCQ-37).

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.

**Response:**

We have removed the original Figure 1 in the newly revised manuscript because its role is not matter to the two main research questions of Web-CAT application and skewness analysis feasibility.

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.

**Response:**

In response to the reviewer’s comment, we have conversely changed the original figure 1 and 2 to let the whole paper in accordance with the structured organization in order of CAT presentation, skewness over years in comparison of group measures (Figure 2) and group on a specific item (Figure 3). We have properly interpreted Figures of skewness analysis applied to distinguish differences between groups.
according to the skewness bands overlaid or departed apart.

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?).

Response:
In this paragraph, we focused readers’ attention on the interpretation of CAT report in Table 3. As suggested by the reviewer, we have reorganized the manuscript structure for ease of understanding.

In the step 2 and step 3 regarding probability (related to person measure and item difficulty) and Outfit MNSQ indicator used for detecting aberrant responses from examinees, those are very important and extremely difficult for regular readers to understand. All of those are included in the Additional file 1 for readers to practice.

In discussion session of Application of CAT, we recommended readers who are interested in the iteration of person estimation and item calibration (step 2) and the computation of Outfit and Infit statistics (step 3) can visit the website at http://www.eddata.com/resources/publications/EDS_Rasch_Demo.xls and at http://www.rasch.org/rmt/rmt34e.htm or even to practice the Additional file 1 of Excel routine to expand their knowledge of IRT-based CAT.

As for the considerations that should be concerned regarding “aberrant responses”, we particularly designed the outfit MNSQ greater than 20.0 as a criterion for being discarded in data collection due to it too aberrant from most of respondents to believe their honor responding the CAT assessment.

8) The conclusion that -0.37 could overcome the shortcomings of traditional assessments has to be further discussed.

Response:
Due to space limit, we have removed the issue of cut-off point discussed in the original manuscript and focused more attention on the two research questions of Web-CAT application and skewness analysis feasibility.

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.

Response:
As suggested by the reviewer, some detailed discussions of skewness analysis have been raised in the session of “Problems in application and daily use”. Furthermore, we have addressed the advantage of skewness usage not only for the graphical representation but also for the reason of truth that data points are not perfectly symmetric in reality. It might be problematic to compare results with the t test and ANOVA. Due to space limit, we only recommended readers to referring preceding published paper [24] where some discussions and descriptions were reported. The Additional file 1 may be helpful for readers to know detail about the skewness with item-by-item analysis in an Excel routine.

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.

**Response:**

As suggested by the reviewer, we have programmed a Web-CAT module on Internet at http://www.healthup.org.tw/irt_test4/irt_start.htm. That is one of the features in this study for readers to practice in their wish. We also added Figure 1 to describe the CAT flowchart and procedure designed in this study. In addition, the relevant information regarding CAT design, for instance, stop rule, item selection method and person estimation algorithm has been described in greater detail in session of “3.Web-CAT assessment” in Methods, “3.Web-CAT performance” in Results and “(3)Polytomous CAT module developed in this study” in Discussions. Importantly, An Additional file 1 is an Excel module in which we introduced something key terms regarding their calculation such as probability, variance ,outfit MNSQ, Newton-Raphson Method, information and SE in CAT routine.