Reviewer's report

Title: An assessment of functioning and non-functioning distractors in multiple-choice questions

Version: 1 Date: 1 April 2009

Reviewer: Si Mui Sim

Reviewer's report:

Major compulsory revisions:
1. The research questions are well defined, the reporting adheres to the accepted standards and the writing is clear.
2. The methods are generally well described and appropriate. However, I have a few questions and comments here:
   a. The number of examinees per test for 6 out of the 7 tests sampled was <100 (range 73-75). I think this is too small a sample to give an accurate estimation of discrimination power (D) of an item option, when using only the upper 27% and lower 27% (i.e. only 54% of the total number) of examinees for calculating the D value. Have you tried using the discrimination coefficient instead of discrimination index? The advantage of using the discrimination coefficient instead of discrimination index is that with the first method each and every person evaluated is taken into account, while with the second, only 54% of the total number of examinees are considered.
   b. What was the content of the test papers examined in this study? Were they discipline-based, multidisciplinary or fully integrated tests? This may have a bearing on the predictability of a test item for the whole test performance.
3. Although the data demonstrated that in these tests there was a low frequency of test items with >2 functioning distracters per item, it's perhaps over simplistic to draw the conclusion that “ideally multiple choice item should consist of three options: one correct answer and two plausible distractors.” I read in here the assumption that by reducing the number of options per MCQ item from 4 to 3, there will naturally be an equivalent reduction in the proportion of non-functioning distractors. This is not necessarily true as there is no guarantee that an otherwise untrained/unskilled examiner will naturally construct a test item with 2 effective/functioning distractors just because of a reduction in the number of options needed per test item.
4. The decrease from 4-options to 3-options per test time appeared to selectively have favoured the low ability examinees without helping the high ability examinees, albeit the change was small and insignificant statistically. Nevertheless the message is still clear that reducing the number of options helps the lower ability students to pass the examination. It is believed that larger number of options gives the lower ability students more opportunity to show what they don’t know. The question is whether in a test, we should only look for what
examinees know, or we should also be interested in what they don’t know!

Minor essential revisions:
Table 2:
1. Please check the value in the parentheses of “Total” column (23.6) for distractors with both frequency <5% and discrimination >0. It looks larger than expected.

**Level of interest:** An article of limited interest

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**
'I declare that I have no competing interests'