Reviewer's report

Title: Using item response theory to explore the psychometric properties of extended matching questions examination in undergraduate medical education

Version: 2 Date: 17 June 2004

Reviewer: Tom Bramley

Reviewer's report:

General
The authors have well addressed all the minor essential revisions and the discretionary revisions.

The only outstanding issue remains the DIF analysis. It is still my view that, in the context of an undergraduate medical examination, it would be preferable to demonstrate the item difficulty invariance necessary to build an item bank from which tests with equivalent measurement properties from session to session could be constructed, rather than to state (correctly) that it follows from the model.

However, I accept that one of the authors’ purposes in the paper is to demonstrate the technique of analysing DIF within an ANOVA framework, and that this is a valid approach.

I would therefore suggest just a minor rewording to make it clear that differences in item functioning between the three arbitrarily defined samples who happened to take the test on different occasions, if truly significant, must really indicate differences in the learning situations encountered by the students taking the test in each term - (e.g. different lecturers?), and not any intrinsic property of the offending items.

-----------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

-----------------------------------------------------------------------------------

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
Bottom of Page 8 - the word ‘patients’ has crept in to replace students!
Page 12 2nd para - ‘and other’ should be ‘another’.
Page 16 2nd para - suggest ‘the same incorrect response is selected more frequently than the correct response’.

-----------------------------------------------------------------------------------

Discretionary Revisions (which the author can choose to ignore)

-----------------------------------------------------------------------------------

What next?: Accept after discretionary revisions

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

Statistical review: No

Declaration of competing interests: None