Author's response to reviews

Title: Assessment of higher order cognitive skills in undergraduate education: Modified Essay or Multiple Choice Questions?: Research Paper

Authors:

Edward J Palmer (edward.palmer@adelaide.edu.au)
Peter G Devitt (peter.devitt@adelaide.edu.au)

Version: 2 Date: 25 July 2007

Author's response to reviews: see over
Dear Editors,

Thank you for the opportunity to revise the manuscript MS: 3175165813840472 - The Modified Essay Question: does it still have an effective role in assessment? Research Paper

The reviewers’ comments have all been helpful and we have made the following changes to the manuscript to address their concerns. We have amended the title to acknowledge their comments.

Reviewer 1

- We have added to the abstract to improve its ability to inform readers and paginated the document

- In the introduction, we have better defined the research question and what we call ‘effectiveness’. We have expanded the methodology to clarify the explanation of how the tests were assessed and have discussed the importance of validity of assessments.

- We thank the reviewer for the Rabinowitz and Downing references. We have taken notice of their findings and incorporated them into our manuscript.

- As opposed to correlation, we have used the kappa statistic to compare the inter-rater agreement. The ratings show good agreement between raters.

- We agree that the conclusion was overstated and have modified it accordingly. We have considered the other forms of assessment and indicated that there may be alternatives to running standard MCQs or MEQs.

- We have provided a new paragraph in the discussion comparing pros and cons of different assessment formats. To address concerns about the tables and figures, we have provided better description and legends on the figure. We have retained table 1 but have also summarised the content into the text, to provide better readability. We have provided a clearer legend for Table 2.

Reviewer 2

- The authors recognise the value of being able to further generalise the results of this study. We have added information in the results and discussion section addressing the weaknesses in the study and its generalizability. The authors of the questions were not available to provide additional analysis depth.

- We were intrigued by the work of Berner et al as recommended by the reviewer, and found it likely to be of interest to readers of the manuscript. We have included discussion in this area.

- We have provided examples of the coding of questions from the assessments, as recommended, including an explanation of the reasoning behind the classification. We were unable to provide an MEQ with three differently coded stages but have instead chosen 2 stages from one problem and 1 from another.
Reviewer 3

- We believe that our paper highlights the difficulty that academic staff have writing test items. Despite the literature supporting the use of one format over another, in practice designing assessments is extremely challenging. We believe that this observation alone is very important and worthy of further discussion. Our paper addresses this issue.

- Processes and skills involved in assessment design need to be of a high order to provide effective assessment. It is therefore appropriate to examine the assessments of one institution and make comparisons between test items. We agree with the reviewers comments that designing MEQs with little instruction and single person review can be one of the major weaknesses of MEQ and MCQ construction but rather than being a confounding factor, it is in reality one of the important points of this paper. Readers may carry out similar evaluations in their own areas and be able to support or refute the conclusions of the authors. We appreciate that we may not have clarified the limitations of this study and have addressed this in the discussion.

- The authors have 20 combined years of assessment experience. One of the authors is the Chair of the MCQ Board of Examiners of the Australian Medical Council. We leave the decision on whether to publish this information in the hands of the editors and are happy to include this information should it be required.

- We have elaborated further on the evaluation process to clarify the process of grading. We also agree that the paper by Haladyna is very relevant to the topic at hand and have included it as a reference.

- There were 145 stages over 30 MEQs, thus making the average number of stages per MEQ similar over all assessments. This was not clear in the manuscript and has been corrected. We have also provided information on the type of MCQ (1 from 5) used in the assessment.

- We have modified the conclusion and discussion to better suit the evidence of the study as recommended.

- We have modified the title, but have attempted not to make it over long. We believe it now better reflects the content of the paper.