Reviewer’s report

Title: Modification of an OSCE format to enhance patient continuity in a high-stakes assessment of clinical performance

Version: 2 Date: 1 April 2011

Reviewer: Matt Homer

Reviewer’s report:

Thanks for the changes you have made. I think that the article is very clear now. I just have a few minor suggestions that you might wish to address, mostly to help the reader in terms of clarity. I think most of the suggested changes would improve the manuscript a little.

Discretionary Revisions

Method, 2nd para: 'the modified OSCE contained'. Would not 'consisted of' be better here than 'contained'? Currently, it might be interpreted that there were other stations etc.

Study sites: I notice that no account was made of 'site' in the analysis - a brief note on this issue might be worthwhile.

Outcome measures: Since there are two raters, it is not entirely clear how a student passes - presumably if their mean rating is greater than 3 (or 3.5?). This could be clarified.

Statistical analysis: The reason for doing the ANOVA is not specifically stated. This would be useful to the reader. Also, the manuscript currently states that 'Pearson correlations between sections were calculated.' It would be more accurate to add 'both within and across scenarios'.

Descriptive statistics: As above, for the ANOVA results, the findings could be interpreted for the reader. Something like 'This means that the participants did not vary significantly in their performance across either section or scenario.' - the link with the negative (=zero) variance components could also be made. Perhaps the ANOVA stuff is actually extraneous?

Variance components analysis: add ‘in the design’ to ‘was unaccounted for’. Also, in the next sentence, why not emphasise that you want high variance component for participant in order to discriminate between them – the whole purpose of the assessment! Finally, following the g-coefficient formula the double subscript term (n-i-s) is not explained.

Inter-rater reliability – there are many intra-class correlation coefficients. It would help the reader to be more specific about this.

Discussion, 2nd para: ‘The consequences of the assessment tool…’. I don’t think
‘consequences’ is the right word here – perhaps ‘efficacy’ or similar?

Discussion, 5th para: re the Halo effect, would it be better to state ‘could be due to a halo effect across sections, within-scenario’?

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.