Reviewer's report

Title: Gender plays no role in student ability to perform on computer-based examinations

Version: 2 Date: 22 June 2006

Reviewer: Christopher Ricketts

Reviewer's report:

General
This is potentially a useful study, as there is a long-held myth that gender does play a role in computer-based testing and, in particular, disadvantages females. It would help those who use computer-aided assessment to have this study published. However, for the paper to have real impact the analysis must be rigorous. I have 2 major concerns and a number of minor issues.

The authors have used Microsoft Excel’s Data Analysis ToolPak to analyse the data. It is impossible to tell in advance whether the following changes will affect the conclusions of the paper, but they will certainly add to the rigour of the analysis.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)
1. First, the authors have been unable to apply the best analysis, because Excel doesn’t have the facilities. There are 3 factors of interest here which may affect performance as represented by score. The first is year group, the second is gender and the third is type of test (paper or online). The analysis should take these into account simultaneously, as simply looking at one factor at a time may mask effects. The authors should take statistical advice on multi-factor ANOVA.
2. Second, the statistical functions in Excel are not of the best and it gets some things wrong. The deficiencies are well known in the statistical community. Any analysis worthy of publication should be based on reputable statistical software (and appropriate statistical analysis as above).

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
3. Tables 1 and 2 are unnecessary and are largely explained in the text.
4. The statistical presentation is largely uninformative as it is based upon copying and pasting the output of Excel’s analysis. Tables of means are usually more informative than the values of test statistics. All tables should be reviewed so that useful information is included. P-values may be better quoted in the text.
5. The authors need to be clear about why they have chosen to use one-tailed or two-tailed tests.
6. A large section of the Discussion (the page before the Conclusions) is not really discussion but largely repeats the findings from Results.
7. The References should be checked for completeness (eg 3) and abbreviations (eg 6, 10, 11 and 12).

Discretionary Revisions (which the author can choose to ignore)

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory 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:
I declare that I have no competing interests.