Reviewer's report

Title: The influence of the cultural climate of the training environment on physicians' self-perception of competence and preparedness for practice.

Version: 5 Date: 30 September 2008

Reviewer: Douglas Curran-Everett

Reviewer's report:

General

The authors have addressed some comments but one question was apparently not addressed. I do have some remaining questions.

Major Compulsory Revisions

1. Power analysis. I am not clear on what makes G*Power unique. On p 8, what outcome was used to estimate the sample size of 40 subjects?

This was my original question, but I am not sure the authors understood the focus of my question, so I will try again. There are other software packages (for example, SAS, PASS) that perform power and sample size analyses. What makes G*Power unique? Why not use one of the other packages?

In order to estimate sample size, among other things, you need to specify the magnitude of a difference you would like to be able to detect. It is not clear from pp 7-8 what outcome (response variable, dependent variable) and magnitude were used to estimate sample size.

2. Statistical analyses. Why were parametric procedures (t test, analysis of variance, regression) used on ordinal data? Were assumptions for proper use of these procedures satisfied? Nonparametric procedures are appropriate in this situation. Descriptive statistics can be reported as medians and interquartile ranges, although with a 4-point Likert scale, it is difficult to see how an estimate of variability will be useful.

I am familiar with the Central Limit Theorem. I teach it in my classes. :)

Parenthetically, 30 observations is not always sufficient for the theoretical distribution of the sample mean to follow a normal distribution; this is not a sufficient criteria to be confident that the theoretical distribution of the sample mean is normally distributed. If a Mann Whitney test produced a similar result, that argues that the parametric analysis is appropriate. The authors must provide explanation and justification for their analysis within the body of the paper.
4. Bonferroni adjustment for multiple comparisons, p 10. Why was the Bonferroni procedure used to control for multiple comparisons? Although it is effective at controlling the overall error rate, its power is quite low.

The advice of an expert is not sufficient justification for the use of Bonferroni. Why did this 'expert' recommend Bonferroni? Some people might consider me at least knowledgable, if not an expert, and I might recommend something other than Bonferroni.

Minor Essential Revisions

2. Questionnaires, p 8, line 15. On what basis were the 4 incorrectly filled-out questionnaires excluded?

Page 8, lines 15-17 do not help me understand how the [now 6] questionnaires were incorrectly filled out. Please provide more specifics in the paper.

3. Percentages, Results and Tables. Is 0.1% all that meaningful? Please report percentages to the nearest integer. The values will be easier to read as integers, and you want your manuscript to be as easy to read as possible.

The Tables are easier to read with rounded percentages, but there remain percentages within the body of the Results that are reported to 0.1%. These also need to be rounded.

4. Results, influence of training on level of preparedness, p 11, lines 12-14. If you want to make a point of potential differences based on year of graduation, age, and training setting, why did you not do a post-hoc power analysis as you did for the primary outcome?

I did not see a response to this comment.

Discretionary Revisions

None

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

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.