Reviewer's report

Title: Attitudes towards statistics of graduate entry medical students: exploring the perception of difficulty and willingness to engage

Version: 1 Date: 23 December 2013

Reviewer: Gill Price

Reviewer's report:

‘Attitudes towards statistics of graduate entry medical students: exploring the perception of difficulty and willingness to engage’

By Ailish Hannigan1,2*, Avril C. Hegarty3, Deirdre McGrath1,2

This article reports results of a survey of students at the start of a postgraduate-entry medical degree about their attitudes towards statistics. It uses a previously published questionnaire, the Survey of Attitudes Towards Statistics (SATS). The idea seems interesting and relevant, but there are problems with analysis, interpretation and presentation of the data. Although there are several aspects addressed in the questionnaire the authors focus on the aspect of perceived difficulty of statistics, and they relate the scores to students’ characteristics and their previous experience of learning quantitative methods.

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

1. Treatment of students who had not previously taken quantitative courses

The authors report that 85% of the respondents in this Survey of Attitudes Towards Statistics (SATS) had previously taken a quantitative course. They do not address the results of the remaining 15%. If these students had never previously been exposed to mathematics or statistics how could they be expected to answer the questions on perceived difficulty of Statistics? Did they provide neutral answers to these questions?

Associations were explored between the scores of questions in the Difficulty factor of the SATS and characteristics, but they could usefully explore the binary variable expressing whether or not the students had previously taken a quantitative subject.

2. Lack of data to support conclusions

There are many places where the data are not presented which are needed to support statements and conclusions. For example, in Table 1, all of the data which are used to relate to scores in the SATS should be presented, for example (but not restricted to) the number of quantitative courses previously taken, perceived ability/performance in maths and the ‘willingness to engage’ variable.
The abstract contains data which are not in the results section, eg. the number (%) of students who said they would choose to take a course in statistics, and the reported expressed feelings of students. The Discussion also refers to data which are not presented under ‘Results’.

Several times an association is reported in the article with only a p-value as evidence, for example in the second paragraph of the Results under ‘Scores on components by demographics’ and elsewhere in the paper. Data expected are at least the effect size (showing direction of effect) with its SE or a 95% CI. Summarised raw data for each comparison group are preferable in addition.

In the paragraph explaining Table 2 (‘Scores on components by demographics’) there is mention of “significant predictors of one or more components” and females scoring significantly lower than males, with and without adjustments, without any data to support this. Effect sizes and SE or 95% CI would be expected.

3. Coding and calculation of answers to the SATS questions

In the Methods section (‘Instrument’ para. 2) the authors say that negatively worded statements were reverse-coded. It is not clear how this worked. Was the reverse coding applied in the possible answers of the questionnaire, or were the answers all kept in the same format and the numbering reversed in coding? Was Cronbach’s alpha calculated on the original or reversed codes?

Interpretation of the data in Table 3 on items of the ‘Difficulty’ component is not clear. It would help to be given interpretations of the individual mean scores which are presented – the interpretation given in the first para. under ‘Factors associated with the perception of difficulty’ needs clarifying. Two of the mean scores seem to have the opposite interpretation in a general sense from the other five, and one wonders whether the means presented are of reverse- or of directly-coded answers.

It is difficult to reconcile the means for items in the ‘difficulty’ component (Table 3) with the overall mean for this component which is reported in Table 2. How was the latter mean calculated? Was there any weighting involved? Could ‘n’ be given for each item in Table 3?

Some justification should be given for re-coding the neutral along with the positive views on ‘willingness to engage’ in the binary variable.

4. Statistical methods

Spearman correlation coefficient: In each of the variables which were related with others using Spearman’s rank correlation, how many categories represented the data? This should be apparent in the Data section of the Methods, or in Table 1, or both. Was Spearman really suitable for these analyses? Would some form of the chi-square test not suit the data better?
How were the data distributed which were analysed as dependent variable using Linear regression?

5. Rationale for focussing on Difficulty

The authors report that there are six components to the SATS. They focus in this paper on one of them. The rationale for this is not very clear.

Why did they focus on Difficulty and not, for example, on Effort (for which there were much higher mean scores)? It seems from Table 2 that the student sample had, on average, more positive-than-neutral scores for Cognitive Competence, Value, Interest and Effort, and more negative-than-neutral scores for Difficulty and Affect. It is not inconsistent to have a perception of something as being difficult and yet to be willing to tackle it as a worthwhile challenge. Why focus so heavily in this paper on the negative aspects of these students’ attitudes towards statistics when there is apparently much that is positive to celebrate and to harness?

6. Timing of survey

What was the rationale for surveying these students at the beginning of their course (especially given that a significant minority of them have no previous experience of quantitative subjects)? Is there a plan to re-survey them after they have learned some statistics in their course?

7. ‘Willingness to Engage’

I have difficulty with the interpretation of students’ indication whether they would have chosen to take a course in statistics or not as being ‘willingness to engage’. The latter seems to be addressed more obviously by the component ‘Effort’ in the SATS. There could be reasons unrelated to willingness to engage behind a student choosing to take a course. One could be whether they feel that they have covered the material sufficiently, or whether they wish to do something different from before, or whether they perceive the course to be relevant or useful to their goals. Why was the component ‘Effort’ not used instead to express ‘willingness to engage’ in the analyses for this article?

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1. In the paragraph explaining Table 2 (‘Scores on components by demographics’) there is mention of interactions, but there is no data to evidence this. From the example following the statement it sounds more like an association than an interaction.

2. Were amendments made to the SATS questionnaire to make it more applicable and intelligible to Irish students? Some of the questions appear to refer to American systems and contexts.
3. Distinguishing statistics from mathematics
The authors suggest that students should be assured that statistics as a subject is not the same as mathematics (eg. in 'Background', para. 3). Was there any attempt to ask separately about students’ previous experience of statistics as distinct from their experience of mathematics?

4. Over-Interpretation of presented data.
Interpretations of the scores (eg. in the abstract) as being ‘reasonably positive’ seems to be stronger than justified by the data in their context. Most of the mean scores were closer to 4 (the neutral score) than to 0 or to 7. A value of 3.7 for ‘Affect’ is only 0.3 of a point from the neutral position, so ‘slightly less positive’ might be a better description than ‘less positive’.

5. Some references seem needed for the first sentence in the third para. of the ‘Background’ (starting with “the relationship between attitudes….”) There is an error in the year of the article by Sackett et al in the ‘References’

6. The hyperlink inserted in the text as reference to the SATS questionnaire should be formatted in the usual way for references.