Author's response to reviews

Title: Determining the quality of educational climate across multiple undergraduate teaching sites using the DREEM measure.

Authors:

Rajesh Varma (r.varma@bham.ac.uk)
Ekta Tiyagi (sanjueku@yahoo.co.in)
Janesh K Gupta (j.k.gupta@bham.ac.uk)

Version: 3 Date: 21 January 2005

Author's response to reviews: see over
Dear Sir or Madam:

We would be extremely grateful if you would consider the revised article for publication in your prestigious electronic journal. We have addressed every single comment from all three referees and modified our paper accordingly. We feel that the paper has been greatly enhanced by incorporating the referee’s recommendations and we would be grateful if you would favourably consider the paper for publication.

We endorse the view shared by all three referees that this manuscript is of great importance in the field. Furthermore, we especially support the view by referee 1 (Sue Roff) who stated of the originally submitted manuscript: ‘This is a well-written report…’ that was ‘…worthy of publication’. We are flattered by her compliments as she was the original developer of the DREEM questionnaire and this makes her remarks of particular high importance.

We list a detailed breakdown point-by-point of the referees comments with this letter.

Conflicts of interest
None declared.

Many thanks

Rajesh Varma r.varma@bham.ac.uk
Ekta Tiyagi sanjueku@yahoo.co.in
Janesh K Gupta j.k.gupta@bham.ac.uk
Replies to referees

Referee 1 (Sue Roff)
This is a well-written report, albeit on data from 2000.
It would be interesting to know the number of respondents from each site.
The number of respondents from each site is listed in table 2 and therefore not repeated.

Also how they managed to get 100% response rate.
We achieved 100% response rate as this questionnaire was distributed and collected at the final MB ChB undergraduate exam. We have added this to the methods text.

Are the 206 respondents the full class? But worthy of publication.
206 was the full class and we have inserted this in to the methods section

Referee 2 (Ian Symonds)

Background

Please clarify if there were any significant differences in way the curriculum was delivered between centres.
There are no significant differences in the way curriculum is delivered between centres, and this has been inserted in to this section

Please clarify that the null hypothesis being tested here was that ‘there would be no difference in the learning environment between different centres’ and the reasons, if any, for asking the question.
The null hypothesis being tested is that there is no statistically significant difference in the learning environment between centres-this has been inserted into this section

We have stated the reasons for the study-this has been inserted into this section: ‘Based on previous student feedback reporting differences in educational experiences, together with our concerns relating to the impact of new curriculum changes, we wished to objectively assess whether the educational environment perceived by students varied at different teaching hospital centres, and whether the environment was at an acceptable standard.’

Method

Please state when during the student attachment the DREEM was administered.
We have stated that the DREEM measure was administered at the end of the eight-week course module; this has been clarified in the methods section.

If this was at the end of the year there may be a need to discuss effect of recall bias.
If at the end of each block did the authors look to see if the overall ratings changed as the year progressed?
We have discussed this issue in the discussion section stating that: ‘By performing the
DREEM survey immediately at the end of the obstetrics and gynaecology module, and emphasising reporting only the last eight weeks experience, we believe this maximised the chance that the DREEM measure assessed only the recent hospital teaching site and minimised any recall bias.

Results

If the Tukey-Kramer comparison is superior why present the results of the T test and ANOVA as well?

We entirely agree with the referee that statistical analysis is a complex and potentially flawed issue with Likert like questionnaires. In the discussion we cite an expert paper that highlights this controversy, and we have deliberately been non-judgmental and used all three statistical tests to present results. Because all the results were non-statistically significant we feel further discussion and elaboration is beyond the scope of this paper and would be immaterial to its aims and conclusions.

Does this mean that the authors consider that the lower scores for learning and teaching at some units do not need to be addressed?

The lower scores for Academic self perception (61%) and Social perception (68%) domains were not statistically significant. We have restated this in the discussion opening paragraph as: ‘The two lowest scoring contributory domains, academic self-perception (61%) and social self perceptions (68%), were not statistically significantly different from the other three DREEM domains or overall mean DREEM score.’ Consequently, we did not individually address these domains as a consequence of the survey as the referee implied.

Discussion

In the discussion the authors state that the overall DREEM score was less than they expected, despite being greater than previous studies. Would the authors please expand on the reasons for this and provide a one-sentence summary of the findings from previous surveys for comparison.

We have discussed the examples of other DREEM scores and reasons for our higher DREEM score by inserting two sentences in this section

Could the authors give specific examples of how the results of this study have been used to improve the delivery of medical education?

We are unable to provide specific examples. However we do state in the discussion last paragraph how we provided feedback to our tutors following the DREEM survey to encourage them to continue with maintaining high quality educational environments.

Discretionary revisions

Should the legend for Table 3 read ‘no statistically significant differences.’? (Rather than comparisons)?

We have altered the legend to read comparisons showed no statistically significant differences as recommended by referee
Can the authors provide any evidence from the literature linking learning styles with student perceptions of the learning environment?
We had not considered this issue, and we are pleased that the referee commented on this. We have found no direct evidence from literature linking learning styles with student perceptions of learning environment, but we did find a report that suggested that greater self-motivation and attitude to learning might bias the DREEM evaluation. We have inserted a sentence citing the Sobral reference to elaborate on this.

It seems unlikely that the topic areas in the DREEM are independent variables – it might be worth commenting on this in the discussion.
We agree with this point. We have inserting a separate sentence in this section and connected it to the above point.

A key finding is the low academic self-perception – do the authors have any data from the formative true/false papers on confidence of answers that might correlate with the DREEM results. We do not understand this question, and are concerned the referee may have misinterpreted the manuscript. The low academic self-perception finding was not statistically significant and is therefore not a key finding. Our aim in this study was not to correlate the DREEM survey with any MCQ formative web based assessment.

It would be helpful to include examples of questions from the DREEM to illustrate each of the topics listed in Table 2.
We have not included these questions as we have cited the original DREEM survey study and all these questions are accessible within this study. We feel that if we list a sample of questions this may unnecessarily prolong the manuscript and merely duplicate the original manuscript’s work.

Referee 3. Hettie Till

Major revisions
1. I have read this paper quite a few times now and have serious problems understanding just exactly what the authors were trying to achieve as there is no clear research question…However, again this statement is not elaborated, i.e. the usefulness of the DREEM for what purpose?
We have re-written the background, introduction and our curriculum sections as recommended. We have deleted the confusing introduction statement of learning styles and clarified our aims and reasons for the study.

2. I find the introductory part ("Background" and "Our Curriculum") rather cryptic and wandering…. Does this mean that only the clinical rotations take place in the 8 different hospitals? There is no explanation of whether all the other modules that make up the final year happens in the same manner and whether a group of students stay in the eight teaching hospital sites for the entire year.
The Methods section is only 5 lines long and does not explain what is being measured. Were students told, for instance, that only the 8 weeks of Obstetrics and Gynaecology must be kept in mind whilst completing the Inventory? We agree with the referee completely. We have clarified how the clinical rotations and lectures take place in the Background and Curriculum sections as recommended. We have expanded our methods and introduction section to fully state methods.

4 Results.
4. This section should be looked at by a statistician. The authors "prove" and "disprove" statistical differences between the 8 teaching sites, which are apparently based entirely on mean scores for the 5 domains then calculated into percentages. A nice little paper by Susan Jamieson "Likert scales: how to (ab)use them" appeared in Medical Education (2004;38:1212-1218) questioning the use of means and parametric statistical analyses used for ordinal data. We agree with the referee and commend her for identifying the statistical complexities and recommending a reference guide. Consequently, we have stated the potential flaws of means and parametric statistical analyses in our discussion section and stated the reference by Susan Jaemison as requested. Our study, as other groups {Jamieson, 2004 21 /id}, has highlighted the potential flaws in using means and parametric statistical tests on ordinal data from Likert scales. As there is no established guidance, we have adopted to show the Student’s test alongside ANOVA and Tukey-Kramer tests to fulfil best statistical methodology.

5. In this section also a table of results per item and per teaching site should be given (see for example Al Hazimi, Al-Hyaiani & Roff’s latest paper: Medical Teacher 2004;26(6):570-573). This information could give an indication of exactly which areas of the educational climate are perceived as less than satisfactory and indicate differences between the 8 teaching hospitals in specific areas. However, such a table of mean responses again does not give an indication of skewed or polarized distributions and a calculation of where most students 'agree' or disagree might be a better indication of areas of weakness (see the Jamison paper mentioned above, as well as Till 2004 - referenced as No 7 in this paper) We feel that our table 3 does provide a comprehensive breakdown of each item per teaching site. Furthermore, we have stated the degree of variation between domains in the results section as an alternative measure. This is stated in the results section as: ‘Greatest variation between hospitals occurred in the Students’ Perception of Atmosphere domain, where there were four hospitals beyond the 95% Confidence Limits; this compared to three hospitals beyond 95% Confidence Limits in all other DREEM domains (Table 2).’ We are therefore unsure what this referee is querying. Furthermore, none of the other two referees have commented on missing tables, or inadequate table data, and therefore we have not altered our tables in any other manner.
6. Discussion
This section is not clear and should be re-written. The authors apparently used small differences in percentage values calculated from mean values to make certain Curricular Changes. In the Abstract they state, "This information may be usefully employed to improve our medical education delivery and strategy, and target areas that would not necessarily have been identified without the DREEM tool". In the Discussion they state, "We have used this information to improve our medical education delivery and address identified shortcomings. These targeted areas would not have necessarily been identified without the use of the DREEM tool".

As stated by the previous referee (referee 2 Prof Symonds), we entirely agree that there is slight confusion. We have re-written the discussion to clarify these points. There were no individual DREEM target areas that needed addressing as none of the domains were significantly different (this is also stated in the above comments to the previous referee). The discussion’s last paragraph states: ‘The non-significant differences between the DREEM domains and between hospitals were significant findings. This was conveyed to our tutors based at the various teaching centres as a positive and encouraging result.’ We hope this resolves the potential confusion and misinterpretation of the study’s result and impact on our practice.

Minor
7. Line 9: Capitals required in Dundee Ready Education Environment Measure
8. Line 10: (DREEM) Inventory NOT "tool". (Twice in the same line).
9. Line 12: The five domains not "modalities".
Our Curriculum P4
10. Line 4: Insert "that" between "ensuring" and "the" at the end of the line
Results P5
12. Line 2: Insert "the" between "of" and "DREEM" and 'Inventory" after "DREEM"
Discussion
13. Line 1" Insert "Inventory" after "(DREEM)"
14. Line 4: Delete "score" before "DREEM" P6
15. 4 lines from the end" principal NOT "principle".

We have made all of the above minor changes as the referee recommended.