Author's response to reviews

Title: Mapping medical careers: Questionnaire assessment of career preferences in medical school applicants and final year students

Authors:

K KV Petrides (k.petrides@ioe.ac.uk)
Chris IC McManus (i.mcmanus@ucl.ac.uk)

Version: 2 Date: 19 August 2004

Author's response to reviews: see over
General response.

We are grateful to our reviewers for their careful reading of the manuscript, and for their useful comments.

Reviewer: Linda L Gottfredson

General

The paper provides a good overview and integration of three empirically-based career development theories that are useful for understanding choice of medical careers. This allows the authors to look at medical career choice in a more informative way. The RIASEC-based description of the 6 careers at the end of the paper is very good.

No response required.

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

1. The 3 applicant samples are compared to the 3 samples of final-year students. The latter are considerably smaller than the former, suggesting considerable attrition. Some analysis of loss to followup is required. How do the students remaining at follow-up differ from those no longer in the sample? To what extent is the loss due to leaving the question blank vs. going to a different medical school vs. leaving medical school vs. refusing to participate?

The referee is entirely correct, and on looking back at our paper we realised that two minor but confusing errors had been made, one in which we described the samples incorrectly (only the 1991 cohort was of entrants to medical school, the others were of all applicants, not all of whom were accepted). The other error was that we had not made clear that the INDSCAL analyses were based only on those respondents who had complete data on all of the career preference measures (and the numbers we provided were based on those groups, which in some ways are slightly misleading as to the studies as a whole). We have now entirely rewritten several sections in order to correct these errors and to clarify the data that we have used. We are very embarrassed that they had slipped into the paper, we once again thank the reviewer for picking up that there was a potentially serious anomaly which would confuse the interpretation, and we sincerely apologise for the inconvenience and extra work that this must have caused the reviewer.

We now hope that everything is correct. (and we have rerun the calculations from raw data in order to confirm the validity of the results we report). In addition we have included in the Supplementary Material a table providing the Ns for each group, along with the breakdown of males, females, mature and non-mature students in each group, and the numbers with complete data for all career measures.

The modified passages of text (which we believe are now consistent) now read:

Sample sizes for the medical student studies were 1135, 2032 and 2973 for the students in the 1981, 1986 and 1991 cohorts (and these samples consisted of all applicants in the 1981 and 1986 cohorts, and all entrants in the 1991 cohort), and were 330, 376 and 1437 for the final-year students in the 1981, 1986 and
1991 cohort studies. The INDSCAL analyses were restricted to those subjects for whom complete career information was available; this consisted of 538 applicants and 312 final-year students in the 1981 cohort, 1118 applicants and 301 final-year students in the 1986 cohort, and 1638 entrants and 1437 final-year students in the 1991 cohort. See Supplementary Information for details of the breakdown of samples by sex and maturity.

2. Knowing more about loss to follow up might help readers better assess two particular reports in the paper.

a. The authors say that differences in results (Figures 3-8) for the applicants and final year students are similar, which is true, but they are noticeably different. I wonder whether the differences are due partly to the substantial loss to followup.

Although the numbers are now much more consistent, because we have reported the correct ones, there is still an important issue here. The analysis of table S3 of the Supplementary Information now assesses this issue properly, and it seems clear that there is no difference in the mean values. Given that it is highly unlikely that there would be a difference in the correlational structure, because it is rare to have structural differences in the presence of invariant means.

b. By not discussing sample attrition, the authors' Figure 9 invites inferences about longitudinal changes in preferences from Time 1 (applicants) to Time 2 (final year), but that is unwarranted due to loss to followup. I presume that the question could be addressed in an analysis limited to applicants remaining in the sample till the final year.

We hope this question is to a large extent answered by the analysis described in (a) 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)

1. How were the applicant samples recruited and the final year students located? Prior publications are available, but readers of this article should be given at least a brief account.

We have extended the description of how applicants were sent their questionnaires, and have also added some extra information in the section on the final year follow-up.

The data were collected during three longitudinal studies of medical student selection, the first of which began in the autumn of 1980, looking at students who had applied for entry to medical school in 1981 [46,47], the second began in the autumn of 1985, studying applicants for entry to medical school in 1986 [48,49], and the third began in 1990, studying applicants for entry to medical school in 1991 [50,51]. The 1981 and 1986 cohort studies were restricted to students applying for entry to St. Mary's Hospital Medical School in London, although since applicants had each applied to five or six medical schools, many students entered schools other than St. Mary's. The 1991 cohort study looked at applicants to five different English medical schools, and because each applicant applied to several schools, these applicants represented 70% of all applicants and entrants to UK medical schools in that year. In each survey, applicants were sent questionnaires as soon as possible after UCCA, the central universities admission system, had received their application and entered their names had been onto the computer database. In general this was many weeks or even months before applicants were asked to come for interview, or were sent decisions on whether they had been accepted or rejected. The data are therefore to a great extent properly prospective.
For the present paper, the analysis of the 1981 and 1986 cohort data considers data on all applicants who replied to our questionnaires, whereas the 1991 cohort, which was very much larger, considers only those questionnaire respondents who entered medical school (although we will generally refer to this group as ‘applicants’ since that reflects the time at which the questionnaire was completed). In the supplementary information we present separate information for the entire 1991 cohort which shows that there are unlikely to be response biases, either due to differences between accepted and rejected applicants, or due to not all entrants responding to the final-year questionnaire. Response rates in the 1981, 1986 and 1991 applicants surveys were 85%, 93% and 93% [46,48,50].

Discretionary Revisions (which the author can choose to ignore)

1. How sure are the authors that students in the first year lab class didn't just make up the data for their 12 "subjects"?

   It is a very fair point, and one obviously has to be careful using such data. This is one of the first practical classes carried out by students in which they are responsible for writing the questions in a questionnaire (not the part we are reporting here), but for collecting and analysing the data. We are therefore confident of their motivation. The class has been carried out for five successive years, and in the first year we were concerned about the problem of data fabrication. An effective way of detecting it is to aggregate the entire set of attitude data collected by an individual student and plot the mean and SD across all of the attitudinal data. Although it is relatively easy to create data for which there is a reasonable location and spread, it is difficult to make this coincide with the average values of the data collected by the other students. In the first year we did the class we detected one very clear outlier using this method, and we reported the result back to the students. Since then at the start of the class we have shown that analysis to each year and since then have never had another outlier. We therefore have confidence in our data. It is also the case that even a small percentage of ‘anomalies’ is unlikely to affect the correlational structure of the data.

2. page 12: Where is the "suggestion" that a third dimension may be necessary for the applicants? The stress levels for the third dimension are the same (.112) for both applicants and final year students.

   We agree that the effect is not large. However it is the relative drop in the stress values that determines the appearance of the Scree plot and, consequently, the number of dimensions to be extracted, rather than the absolute values per se. Thus, although the third stress value is the same in both the applicant and the final-year data, the drop is somewhat more pronounced in the former dataset, which creates the impression of a third dimension possibly being present there. Overall we do feel confident that there is no substantial third dimension present in our data. It may be that we were being overly cautious in our analysis, and even taking the possibility into account, but we didn’t wish to be seen as rushing straight to two dimensions on the grounds that that would be most consonant with the Holland typology.

3. It would help to label the axes in the figures.

   This has now been done (see also a comment to the other referee who noted problems with figure 9).

4. typographical error on page 14, middle: 191=1991

   Corrected.
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

Acceptable Quality of written English:

Statistical review: No

Declaration of competing interests: None
Michael M. Goldacre: Reviewer's report:

General -

-----------------------------------------------------------------------------------------------

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

-----------------------------------------------------------------------------------------------

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. Page 12, 'The medical student samples', 2nd paragraph, 1st sentence: needs attention. It may simply be that the word 'and' is missing between 'applicant data...' and '...the combined...'

Yes, an 'and' had slipped and made it near meaningless.

2. Page 13, 'The medical student sample', 3rd paragraph, last sentence: typo in 'attitudes' and perhaps say 'younger' rather than 'non-mature'.

Both have been corrected.

3. Last sentence before Discussion: typo in '1991'.

Corrected.

4. Figure 9: the meaning of the values on the vertical and horizontal axes are not clear.

We have labelled the axes properly, and they relate directly to the comments in the text.

-----------------------------------------------------------------------------------------------

Discretionary Revisions (which the author can choose to ignore)

5. The sample sizes for the six surveys (three cohorts at application and at final year) varied a lot. Are the authors confident that the smallish samples (eg 301 and 312 subjects) are sufficient for the multivariate analyses undertaken?

While a few samples were comparatively small, their absolute size was more than adequate for the INDSCAL analyses. However, it should be noted that the group weights for the mature candidates (male and female) in the 1981 and 1986 final-year data were based on very small samples, ranging between 5 and 17 participants, and may, therefore, be unstable. We have made a note to that effect in the last paragraph of the results section. We have also in table S1 of the Supplementary information give a precise account of the numbers of subjects in subgroups.

And the drop in the 1986 cohort from 1118 subjects at application in the final year looks substantial - is there an explanation for this?
A similar point was raised by the other referee, and we include here the same detailed answer (and apology) that we provided to that other referee, which explains the errors which had crept into the original manuscript.

The referee is entirely correct, and on looking back at our paper we realised that two minor but confusing errors had been made, one in which we described the samples incorrectly (only the 1991 cohort was of entrants to medical school, the others were of all applicants, not all of whom were accepted). The other error was that we had not made clear that the INDSCAL analyses were based only on those respondents who had complete data on all of the career preference measures (and the numbers we provided were based on those groups, which in some ways are slightly misleading as to the studies as a whole). We have now entirely rewritten several sections in order to correct these errors and to clarify the data that we have used. We are very embarrassed that they had slipped into the paper, we once again thank the reviewer for picking up that there was a potentially serious anomaly which would confuse the interpretation, and we sincerely apologise for the inconvenience and extra work that this must have caused the reviewer.

We now hope that everything is correct. (and we have rerun the calculations from raw data in order to confirm the validity of the results we report). In addition we have included in the Supplementary Material a table providing the Ns for each group, along with the breakdown of males, females, mature and non-mature students in each group, and the numbers with complete data for all career measures.

The modified passages of text (which we believe are now consistent) now read:

Sample sizes for the medical student studies were 1135, 2032 and 2973 for the students in the 1981, 1986 and 1991 cohorts (and these samples consisted of all applicants in the 1981 and 1986 cohorts, and all entrants in the 1991 cohort), and were 330, 376 and 1437 for the final-year students in the 1981, 1986 and 1991 cohort studies. The INDSCAL analyses were restricted to those subjects for whom complete career information was available; this consisted of 538 applicants and 312 final-year students in the 1981 cohort, 1118 applicants and 301 final-year students in the 1986 cohort, and 1638 entrants and 1437 final-year students in the 1991 cohort. See Supplementary Information for details of the breakdown of samples by sex and maturity.

6. It is not clear how the answers to questions about career choices are converted into the scores that become the plots on the Figures. Could this be clarified?

We agree that it was not clear to those who have not used INDSCAL. We have therefore added to the Statistical Analysis section a paragraph which summarises the entire process from raw data to the points on the figures.

“The raw data, which were collected on a 5-point Likert scale, were transformed into Euclidean-based dissimilarities for all combinations of career pairs, using the PROXIMITIES program in SPSS. Four different dissimilarity matrices were produced (non-mature males, non-mature females, mature males, and mature females), and these matrices provided the basis for the INDSCAL analysis that involved minimisation, in Euclidean space, of the discrepancies between the career dissimilarities and the corresponding interpoint distances on the map. The loadings of each career on the two extracted dimensions are then plotted onto the figures to provide the maps.”

7. It is not clear what the 'stress levels' mean, in the section on 'The medical student sample', second paragraph, and how they should be interpreted. Could this be clarified?
We have expanded on the concept of stress in order to clarify these issues. In the Statistical Analysis section we have included the following paragraph:

“This dimensionality of MDS/INDSCAL analyses can be assessed, in a manner analogous to that used in factor analysis in which eigenvalues are plotted against factor number. In MDS one plots a measure of ‘stress’ (in effect, the opposite of goodness-of-fit) against the number of dimensions which have been extracted. If too few dimensions have been extracted then the stress is high, the model not accounting adequately for the richness of the data. The optimal number of dimensions is typically indicated by a sudden ‘dog-leg’ in the stress plot”.

In addition, we have expanded the section where we discuss ‘the medical student sample’ to clarify things:

“This dimensionality of the medical student samples was assessed by carrying out a standard multi-dimensional scaling analysis (i.e. MDS, not INDSCAL), separately for the combined applicant data and the combined final year data from the three cohorts. The stress formula attempts to quantify the discrepancies between the fitted distances in the model and the observed dissimilarities among the career ratings, with larger values indicating poorer fit. It is obviously the case that the more dimensions are extracted, the better the fit of the model and, hence, the lower the stress value. However, it is also the case that a greater number of dimensions complicates interpretation and may lead to overfitted and unstable solutions. The stress levels with 1, 2, 3, 4, 5, and 6 dimensions were .352, .174, .112, .077, .059 and .048 for applicants, and .390, .174, .112, .082, .064 and .052 for final year students.....”

We hope this makes the analysis easier to understand.

8. The authors specify 'six prototypical specialities' (Discussion, fourth paragraph). These do not include general practice. Is this because it is not regarded as 'a speciality'? How does general practice fit in to the reasoning?

This is a fair point. In the Discussion section, when providing a pen-portrait of ‘Psychiatry-Artistic’ we had originally stated, “Psychiatrists, and to a lesser extent, General Practitioners”. We have now altered that to include General Practitioners a little more forcefully, since the two specialities are close in the various maps, although General Practice is usually a little to the left.

**Psychiatry – Artistic.** Psychiatrists, and also General Practitioners, have a more artistic approach to medicine, seeing, interpreting and responding imaginatively to a range of medical, social, ethical and other problems. The emphasis in many ways is on the uniqueness of the patient, the ideas that they are expressing, and the psycho-social theories and concepts which are necessary for interpreting the individual.

9. Opening sentence of introduction: it's not so much that medical careers 'end' in this context, but that 'postgraduate training ends with most doctors specialised...'.

*That is indeed a more accurate description, and we have modified the sentence to that effect.*

What next?: Accept after minor essential revisions

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No
Declaration of competing interests: None