Author's response to reviews

Title: Practice effects in medical school entrance testing with the Undergraduate Medicine and Health Sciences Admission Test (UMAT)

Authors:

Ian B Puddey (Ian.Puddey@uwa.edu.au)
Annette Mercer (Annette.Mercer@uwa.edu.au)
David Andrich (David.Andrich@uwa.edu.au)
Irene Styles (IreneStyles@uwa.edu.au)

Version: 5 Date: 10 February 2014

Author's response to reviews: see over
**RESPONSE TO REVIEWERS AND EDITOR**

**Reviewer: Don Munro**

1. Several of the longer sentences in the results and discussion have been punctuated with more commas as requested.

2. The following fuller description of the UMAT test has been incorporated into the background:

“The UMAT is not a test of personality, values or attitudes, nor is it a measure of previous learning in academic domains. It comprises three sections: UMAT-1 - Logical Reasoning and Problem Solving, UMAT-2 - Understanding People and UMAT-3 - Non-verbal Reasoning and is a 3-hour paper-and-pencil test. Each section is assessed using multiple choice questions with 4 or 5 optional responses, of which only one is correct. The three sections in 2012 comprised 48 items over 70 minutes in UMAT-1, 44 items over 55 minutes in UMAT-2 and 44 items over 55 minutes in UMAT-3. In section 1 (UMAT-1) candidates are required to exercise logical reasoning and problem solving skills using both inductive and deductive reasoning, with an emphasis on logical argument in working to a solution. In section 2 (UMAT-2) the emphasis is on assessing empathy and emotional intelligence with candidates required to show an understanding of the thoughts, feelings, behaviour and intentions portrayed within each question. Section 3 (UMAT-3) evaluates a candidate’s non-verbal reasoning skills and aims to measure cognitive ability independent of language ability and specific cultural knowledge.”

3. Justification for the use of percentile ranks - Each year ACER reports candidate results as scores on a separate scale for each of UMAT1, UMAT2 and UMAT3. An Overall score is calculated as the average of these three scores. Over the years 2000 to 2012, the scale has changed several times as the test has developed. As a result scores are not necessarily comparable between years. However the percentile ranks are also reported for each book for each candidate. This gives a measure of the relative standing of a candidate within the cohort. Given the large numbers in each cohort it is likely that there is reasonable comparability in the competence of candidates at specified percentile ranks across cohorts. Hence the use of percentile ranks as the measure in this study. It is considered that
percentile ranks will be at least as accurate as actual scores in this context. This explanation has been incorporated into the Methods section.

4. Reference is now made to UMAT percentile scores rather than UMAT scores throughout the manuscript.

5. The socio-demographic profile by year the UMAT was first sat from 2000-2012 has now been incorporated as Additional File 1. The distributions are reported as valid percent exclusive of missing values. A paragraph has been added to the Results section outlining the changes that have occurred in the socio-demographic profile over this period as follows:

“Of the 135,833 cases where a candidate had sat once or more during the period 2000-2012, 1.1% were aged less than or equal to 16 yr, 35.6% were 17 yr, 43.3% were 18 yr, 8.4% were 19 yr, 9.9% were aged 20 to 30 yr and 1.7% were greater than 30 yr. Females comprised 57.8% of all cases and males 42.2%. English was spoken at home by 69.2%, Asian languages by 27.1%, European languages by 2% and other languages by 1.7%. School of origin was a government school for 47%, an independent school for 32.8%, a Catholic school for 17.6%, TAFE college for 0.4% and Other school for 2.2%. The majority of the population were Australian (86.9%) with 12.1% from New Zealand and 0.9% from other countries. Only 0.4% self-declared as ATSI in origin. For Australian residents (N=118,086), 93.4% were living in Highly Accessible areas, 4.9% in Accessible areas, 1.2% in Moderately Accessible areas, 0.4% in Remote areas and 0.1% in Very Remote areas. The socio-demographic profile by year in which the UMAT was sat is outlined in Additional File 1. It demonstrates an increasing number of 18 and 19 yo sitting the test over time with a decreasing number of those either 17 yo or less or those greater than 20 yo sitting the test. The proportion of females has decreased slightly while that of males has increased. The proportion of those from an Asian language background has almost doubled. The proportion from a government school background has slightly increased while those from an independent or catholic school background has commensurately decreased. The numbers of New Zealand cases increased substantially from 2005 onwards with introduction of the UMAT as a selection tool for the 2 NZ medical schools.”

In addition, Ns have now been incorporated into each of Tables 3 through 7.
6. We have elected to leave the correct B values, rather than artificially reverse the negative and positive values.

**Reviewer: Daniel McLaughlin**

**Major Compulsory Revisions**

1. The following qualification in relation to European language background subjects has now been added to the discussion.

“The single strongest predictor in this study was UMAT percentile score on initial testing. Being male, being younger, being from NZ and being from non-English language backgrounds (although only of borderline significance for those of European language backgrounds) all predicted a greater likelihood of a re-sit even after adjustment for the influence of initial UMAT percentile score.”

2. The suggested modification has been made as follows:

“The pattern of predictor variables for multiple re-sits were largely unchanged in these models, while the relative strength of the odds ratios was increased for each predictor variable together with the relative amount of variance explained by each model.”

**Minor Essential Revisions**

1. The suggested modification has been made as follows:

“The practice effects were further evaluated by dividing the cohort into quartiles of initial test performance and analysing the upper and lower quartiles in relation to number of times the UMAT was sat.”

2. The suggested modification has been made as follows:

“Predictors of the magnitude of any practice effects were evaluated from a number of socio-demographic indices collected on enrolment for the UMAT, including age, gender, postal address, language spoken at home, type of secondary school, country of origin and self-identification as being of Aboriginal or Torres Strait Islander origin (ATSI).”

3 & 4. The 3 figure legends have been modified as requested:
Figure 1: Total UMAT percentile score and percentile score in each of its 3 subsections (± SEM) by number of times candidates re-sat the test between 2000 and 2012. Error bars are absent when they are too small relative to the mean.

Figure 2: Total UMAT percentile score and percentile score in each of its 3 subsections (± SEM) by number of times candidates re-sat the test between 2000 and 2012 for those in the lowest performance quartile in their first test. Error bars are absent when they are too small relative to the mean.

Figure 3: Total UMAT percentile score and percentile score in each of its 3 subsections (± SEM) by number of times candidates re-sat the test between 2000 and 2012 for those in the highest performance quartile in their first test. Error bars are absent when they are too small relative to the mean.

5. The suggested modification has been made as follows:

“Separate models (data not shown) indicated that for those in the lowest quartile of total UMAT percentile score at initial testing, the predicted increment on re-testing was 18.8 percentiles (CI 18.0, 19.7). For those in the lowest quartile of UMAT-1 it was 22.5 percentiles (CI 21.6, 23.4), for those in the lowest quartile of UMAT-2 it was 33.5 percentiles (CI 32.5, 34.6) and for those in the lowest quartile of UMAT-3 it was 35.7 percentiles (CI 34.7, 36.7).”

6. The suggested modification has been made as follows:

“For the UMAT we now report qualitatively comparable results,”

Dictionary Revisions

1. Please see response 2 to reviewer Don Munro.

Reviewer: Jon Dowell

The authors did not have access to information about applications/offers/admissions to medicine.

With respect to the question on young candidates, the numbers of those 16yo or younger was relatively small (N now listed in each of the tables). In some states in Australia it is quite
possible for a 16 year old to be in their final year at school at the time of sitting the test. There may be a few candidates who have undertaken the test as a trial run in Year 11, however ACER takes steps to avoid this happening, so whilst a few may have slipped through the system, this should not have been a widespread problem.

Observations on what the final implications of what the study outcomes might have for equitably dealing with those who have improved their score on re-sitting would be premature. As stated in the discussion we do not know whether the improved score represents an improvement in competence or just an improvement in performance because of getting the easy questions right and/or guessing. Item level analysis of the UMAT is necessary to gain better insight into this and is currently being undertaken by our team.

**Editorial requests:**

We have added the statement “This study was not commissioned or supported financially by UMAT or ACER.” Please note, however, that in light of the observations reported in this paper we have now been given a grant by the UMAT consortium to undertake the item level analysis referred to above.

The requested modifications to the Figures have been made.