

Selection into Medicine and Health Sciences- the Australian Experience: GAMSAT and UMAT

Cecily Aldous

Cecily Aldous, Australian Council for Educational Research.

1. BACKGROUND AND FUNCTION

In 1994 a group of 3 Australian Medical Schools decided to abandon undergraduate medical training in favour of a graduate-entry model, with entry available to graduates of any discipline, not just those who had completed a science degree. This move coincided with significant changes in medical education and the introduction of problem based learning (PBL) curricula across many medical schools. The new curriculum was designed to provide students with early and on-going exposure to patients in a clinical setting, from day one of the course; to substantially reduce the time spent in lecture theatres and in rote learning, and to increase the emphasis on independent and small group learning. The integrated approach to learning means students are encouraged to work on case studies rather than manifestations of disease in isolation. Thus, it became

important to select students with good interpersonal skills and demonstrated ability to solve problems and think critically rather than, as was the case previously, students who had simply achieved high marks in school level science subjects. At the same time, public demand for doctors who were able to communicate effectively and empathise with their patients was growing.

It was decided that admission to the new graduate-entry courses would be based on three eligibility criteria:

- Grade Point Average (GPA) calculated on the bachelors degree
- Performance in a selection test
- Performance at interview

Opening entry to graduates of any discipline meant that the onus was on the selection test to provide evidence that a candidate had mastered concepts in the

basic sciences. The universities were agreed however that the focus should be on application of the knowledge, in tasks requiring reasoning and problem solving skills rather than on simple recall. Higher order skills in comprehension and interpretation of ideas in general contexts are addressed separately, as is communication in writing. GAMSAT thus became a cognitive test in three distinct sections:

- Reasoning in Humanities and Social Sciences
- Written Communication
- Reasoning in Biological and Physical Sciences (chemistry 40%, biology 40%, physics 20%)

GAMSAT provides a broad measure of knowledge and skills across a number of defined and distinguishable content areas. Biology and chemistry is assessed to approximately first year university level, with physics at the level of final year of schooling. The humanities section focuses on the understanding and interpretation of ideas, involving complex verbal processing and conceptual thinking. In the development of the test specifications the model used in MCAT was taken into consideration.

UMAT had slightly different origins in that it was first developed by ACER for a single university, in 1991, with additional

universities joining in over time until by 2002 it was used by all Australian undergraduate medical schools with only one exception, by a number of health science faculties and by one New Zealand medical school. It aims to provide a measure of aptitude for medical study as evidenced by general non-curriculum based reasoning and problem solving ability and interpersonal skills. It is designed to assess cognitive abilities and skills not directly demonstrated in academic results and is used in conjunction with school results and interview in the selection process.

2. MANAGEMENT STRUCTURE

The graduate medical schools formed a Consortium for the purposes of coordinating the selection process and ACER was contracted to work with them to develop and administer the testing program. Experience with both GAMSAT and UMAT has shown that there are substantial advantages in the testing program being managed by a non-university body, one practised in the development and administration of large-scale testing programs. A further advantage in the current competitive university environment is that, in the absence of competing interests, a non-university manager is well positioned to interact with

all group members and act on behalf of each equally. The neutrality of ACER has been very useful at times.

Both GAMSAT and UMAT programs are managed fully by ACER, with reference to a number of key committees set up to enable effective liaison between ACER and all Consortium member universities. These committees oversee policy development and change, the financial structure of the programs and technical issues, as well as providing a forum for wider discussion on subjects of mutual interest.

One of the early outcomes of this structure in the case of GAMSAT was the establishment by ACER of the Graduate Medical Admissions Centre which processes all applications for admission to the graduate-entry programs at the Consortium universities. A more recent development has been the establishment of the GAMSAT UK program which is currently used by 3 British medical schools and a veterinary science school in Ireland. The GAMSAT Consortium now numbers five medical schools and one school of dentistry.

3. TEST DEVELOPMENT

Detailed test specifications are established and these are operationalised into test questions (items) by teams of ACER test

writers expert in their subject areas and experienced in writing multiple-choice questions. Items are developed to rigorous professional and technical standards. In the case of GAMSAT, faculty specialists from each of the participating universities are also involved in the review of stimulus material. All items in development must pass detailed critiquing, trial testing, statistical analysis and final review. This process results in any items that are shown to perform poorly on any of a number of criteria being discarded. Scrutiny at this stage typically leads to around one third of trialled questions being eliminated as less than optimal. Only items with good performance statistics from the trial testing are used in a final test form. The content, style, duration and sequencing of the tests are determined to ensure that the testing program is relevant, fair, valid and reliable. Items are carefully scrutinised in order to minimise gender, ethnic or religious bias and to ensure the tests are culturally fair. Once a testing program is established, the preferred means of trial testing new items is the in-test trialling method.

More extensive analysis is performed on the response data from the live test. At this point it is possible to match item statistics with a range of demographic data collected from candidates on the registration form. Even at this stage any item appearing to

favour or discriminate against a sub-group in the candidate population can be eliminated from the scoring. The procedures used for the analysis of items and scoring are based on the Rasch measurement model. Among other advantages, use of this model allows for sample independent calibration of items, allowing items to be positioned on a constant difficulty continuum. This can then be matched with the abilities of the candidates each year. It is beyond the scope of this presentation to explore further the psychometric analysis involved in GAMSAT or UMAT.

4. ADMINISTRATIVE STRUCTURE

The testing programs operate on an annual cycle, with one test date per year. GAMSAT in Australia, with approximately 2600 candidates a year, and UMAT with 8500 each require a full time office with one dedicated Project Officer working closely with the Project Director. Administrative infrastructure is established in the first year of operation and processes are continually monitored and enhanced in subsequent testing cycles. An initial policy decision is taken on the level of candidate interaction/response to be supported. Both GAMSAT and UMAT require a very high level of direct candidate contact. However,

it is not the place of the test management agency to provide advice on issues related to university admission. Thus it is important that all staff are clear on the delineation between questions relating to the testing program and those that should properly be referred on to the participating universities.

Key structures to be put in place are a candidate database and registration system. The market environment of each program is analysed to evaluate the security requirements, both in terms of candidate identification and test administration. High stakes programs such as GAMSAT and UMAT where the pressure to succeed is great, and where consequently the rewards of undermining the system could be substantial, necessitate a high level of vigilance. There must be several checks on candidate identity and the integrity of the test must be guarded with great care. This is a costly exercise. The first step in a security system is allocating each candidate a unique identification number to be matched with their records at all stages of the process. Decisions must be made on what information it is appropriate to collect from candidates, how this will be used, how long it will be stored and what if any are the implications under privacy legislation. Some programs may require candidates to submit photographic identification. In GAMSAT a

passport photograph must be supplied with the registration form. This is stored in the database and reproduced on the candidate Admission Ticket which is presented on the test day, together with a specified independent photo-bearing identity document such as passport or drivers licence. Candidates presenting without these documents will be refused entry to the test.

Crucial to the registration system is the Candidate Information Booklet, containing the machine-readable registration form. Information booklets are printed in large quantities and contain all the information necessary to enable prospective candidates to evaluate the suitability of the test for their aspirations and to register successfully. Some CIBs also contain a small number of sample items. The more thorough the CIB, the fewer the phone calls to the program office, in theory at least. A web-based downloadable version of the CIB is also useful. GAMSAT and UMAT both offer paper-based and online registration facilities.

Additional publicity for the programs is provided by a poster and newspaper advertising. Widespread advertising is essential, particularly in the early transition years, to avoid complaints, late requests for registration, etc. All promotional material should display the registration deadline and test date pro-

minently and these dates must be strictly adhered to by the program office. CIBs are provided in bulk for distribution by participating universities and sent out to individuals and schools on request throughout the six months preceding the test.

Registered candidates are issued with a personalised Admission Ticket three weeks before the test. This informs them of the exact reporting address and is a key component of the identification procedures on the test day. A specified number of weeks after the test, Statement of Results slips are mailed to candidates and an electronic file of results sent to participating universities.

5. TEST DELIVERY

Venues suitable for large-scale test administration are usually in short supply, and can be very expensive to hire. They need to be booked a long way ahead of the test date. It is helpful if, where possible, staff selected to supervise are already familiar with the venue. Suitable supervisory teams can be recruited from Examinations Departments of universities or professional test delivery companies. The Supervisors Manual developed by ACER is the key document in this process, designed

to ensure standardised test conditions are followed at every centre, and covers in full detail all aspects of tasks relating to:

- receipt and storage of test materials
- planning for the test
- procedures on the test day
- administering the test
- post-test procedures

Instructions include advice on how to deal with difficult situations such as attempted cheating, candidate illness, disruptive candidates and missing materials, as well as general information on staff/candidate ratios, room set up, etc. All staff are required to sign a confidentiality agreement.

ACER, via the candidate database, allocates candidates to test centres (nominated by candidates on the registration form) and produces an Attendance Roll for each centre.

6. TEST RESULTS AND REPORTING

Completed test papers are returned to ACER where the machine readable optical mark read sheets are scanned, cleaned and edited before being passed to the project psychometric team for further analysis and score generation. Written scripts are batched for subjective marking and raters

trained in use of the criterion-referenced marking protocol and guidelines. All GAMSAT essays are triple marked, with a fourth marking in the case of discrepant scores. These scores are then combined and matched with the multiple choice scores, equating and scaling undertaken and final results imported to the candidate database. Use of modern measurement theory allows the performance of the candidates and the difficulty of the items to be plotted on the same interval scale. Thus the scales for each section of the test reflect the difficulty of items contained in that section and the abilities of the candidates. Candidate raw scores are transformed onto the scale without loss of the original rank order.

Policy must be established on the form of results to be provided to candidates and user universities. The aim with GAMSAT and UMAT has been to enable the greatest flexibility for universities to use the scores in the way that suits them best. Thus, GAMSAT provides four scores; one for each test section and an Overall Score which is a weighted average of the three Section Scores. Each of the four scores is expressed on a scale with a range of 0 to 100. There is a scale for each section rather than a common scale because each section measures a different attribute or skill. User universities are free to set their own cut-off scores from year to year, with some

specifying minimum Section Scores, others a minimum Overall Score and others a combination of the two. Since scores in both programs are valid for two years, a statistical equating process is used to ensure score comparability over time. In a situation where a candidate has scores from two consecutive years, a policy decision should be made on whether there is to be any restriction on their choice of which years results they use in an application.

In the Australian context, taking GAMSAT or UMAT and applying for admission to a medical course are separate events, with candidates deciding whether or not to apply after they have received their test results. However, in UK where GAMSAT is used by a group of three medical schools they have chosen to combine the two processes, so that candidates pay for both test registration and consideration as applicants in one move. Thus, in UK every GAMSAT candidate is also an applicant, whereas in Australia about half of the candidate population in a given year go on to apply for admission in that cycle, and pay an additional fee.

Once test results have been reported, psychometric staff first analyse and report to the test development teams on the trial items (GAMSAT is produced in multiple forms with a small number of trial items in each). The final stage in the cycle is then

the preparation of the annual technical report. This report is for limited circulation among Consortium universities and provides detailed analysis and information on the performance of the test and the candidate population in the given year.

7. TEST PREPARATION

Test preparation and how it is obtained can be a contentious issue for the test development agency. The Australian experience with GAMSAT is probably broadly similar to that of the US in this area. Most candidates would clearly benefit from some targeted preparation or revision before sitting the test, particularly in the sciences, and to this end ACER has developed a number of publications containing sample material (generally items used in a live test and subsequently retired). These publications find a ready market among candidates and their sale provides a valuable financial support to the program. In addition, prospective candidates are advised that more detailed preparation can be obtained through a bridging course, tutor or appropriate text book.

GAMSAT is one of a small number of large scale testing programs in Australia that is also very high stakes. In a situation where the outcome of a single test performance

has the ability to affect the future career of the test taker to a significant degree, commercial coaching courses soon began to flourish. Some of the resulting problems include the development of a belief among candidates that participation in a commercial course is essential to good performance in the test; confusion about what is officially endorsed preparation and what is not; undermining of the authority of the test development and administration agency in matters of advice to candidates. Of major concern to the universities is the issue of equity: the private courses are usually quite expensive and also only

available to city students. Thus if they do confer benefit, it is to urban and affluent students. A number of these commercial operators are expanding their business to offer interview preparation also. Many of them survive by soliciting information from candidates who have recently taken the test, one even being on record as asking students to promise to remember and report back one test question each after the test administration. They are often guilty of misleading advertising. Monitoring their activities and seeking legal advice as necessary is an added cost to the program.