

General Form1April 10, 1996

Development of a Selection Test for Graduate-Entry Medicine

Cecily Aldous

Australian Council for Educational Research, Melbourne, Vic.

BACKGROUND

During 1992, three Australian medical schools (at Flinders University and the Universities of Queensland and Sydney) independently decided to replace their traditional six-year undergraduate medical courses with four-year graduate-entry programs. The programs are open to graduates of any discipline who can demonstrate the skills deemed necessary for completion of the course. The three medical schools subsequently formed a consortium for cooperation in the development and promotion of the graduate-entry medical programs (GEMPs) and have engaged ACER to develop and administer the selection testing and to establish the Graduate Medical Admissions Centre through which applicants to the three schools will be screened.

The development is occurring at a time of intense international scrutiny of medical education throughout the English-speaking world, with change occurring in many countries. America has long had a system of graduate medical training and its admissions test, MCAT, has been running for many years. It is widely agreed that, with the rapid expansion of the medical field, the traditional curriculum is becoming overloaded. The new GEMPs seek to introduce skills of self-directed learning which are seen by many as essential for a lifetime of practice in an increasingly complex profession. There is also an increasing recognition of the importance of effective communication with patients and colleagues. Thus teamwork and an awareness of community concerns will be stressed in the new programs.

Whereas medical training is traditionally based on progressive education in the body systems, the new curricula are problem-based. Students learn medicine in a framework similar to that in which doctors practise. Small groups work cooperatively, with a tutor or at the bedside, focusing on a carefully planned series of patient-centred problems, each designed to highlight a particular set of principles and issues in health and disease. The problem-based framework encourages integration of the basic medical sciences with social perspectives and practical aspects of diagnosis and treatment. Wide-ranging exploration of relevant issues in each problem ensures that appropriate emphasis is given to population and public health issues, ethics, behavioural science and the use of modern medical information systems.

The three schools also had major concerns that existing methods of selecting medical students directly from school - based largely on academic criteria - were unsatisfactory. The immaturity of school-leaver applicants meant that many lacked both knowledge of the profession and insight into their own personal characteristics. It is well known that students achieving high marks in science subjects are frequently encouraged to enter medicine regardless of any vocation, simply because it is prestigious. The medical schools recognised both the desirability of selecting on the basis of demonstrated performance at tertiary level rather than future promise, and of broadening the basis for admission. Gifted students of physics and chemistry at Year 12 level do not necessarily go on to develop the skills in

interpersonal communication and empathy that make for a good practising doctor. It is hoped that the combination of selecting from a more mature pool of applicants and testing them, not just for ability in the sciences, but also for skills in written communication, critical thinking and problem solving across a wide range of subject areas will provide a more balanced cohort of medical graduates.

Selection for the four-year graduate-entry programs will be based on three major criteria: academic performance in a first degree; scores on a test designed to measure basic knowledge and skills appropriate for entry to medical school; and a semi-structured objective interview.

In this paper I will outline the development of the Graduate Australian Medical School Admissions Test (GAMSAT) which has been undertaken by the Consortium of Graduate Australian Medical Schools and the Australian Council for Educational Research (ACER) and comment on some aspects of the first test held in February 1995.

GAMSAT

GAMSAT seeks to evaluate the nature and extent of abilities and skills gained through prior experience and learning, including the mastery and use of concepts in basic science, as well as the acquisition of more general skills in problem solving, critical thinking and writing. The test is divided into three sections designed to assess performance in the areas of:

- IReasoning in Humanities & Social Sciences
- IIWritten Communication
- IIIReasoning in Biological & Physical Sciences

The test takes 5 1/2 hours, with Sections I and II in the morning and Section III in the afternoon. Sections I and III consist of 75 and 110 items respectively, and are in multiple choice format. Section II

consists of two thirty minute writing tasks. Candidates receive a separate score for each of the three sections, as well as an overall GAMSAT score.

Section I tests skills in the interpretation and understanding of ideas in social and cultural contexts. Stimulus materials include passages of personal, imaginative, expository and argumentative writing. Some units may present ideas and information in visual and tabular form. Materials deal with a range of academic and public issues. Questions in this section demand varying degrees of complex verbal processing and conceptual thinking, logical and plausible reasoning and objective and subjective thinking. (skills that emphasise understanding - recognition of explicit and implicit meanings, involving interrelating, elaborating and extending concepts and ideas, discriminating and judging)

In many ways Section II is the most straightforward. It is a test of generative thinking, the first task demanding expository and argumentative writing based on a current affairs type issue while the second deals with more personal and social issues and requires a more reflective and discursive style of writing. This test is assessed on two criteria. These criteria address the quality of the thinking about a topic and the control of language demonstrated in its development.

In addition to testing problem solving within a scientific context, Section III examines the recall and understanding of basic science concepts. These skills include the ability to identify knowledge in

new contexts, to translate knowledge from one symbolic form to another, to estimate measurements and to recognise limits in accuracy, to formulate hypotheses, extrapolate and interpolate, to formulate generalisations in the light of given relationships, to analyse data, make comparisons and follow a line of reasoning.

There is an inevitable tension in any test incorporating 'hard' science and the less tangible skills needed in the humanities and social sciences. GAMSAT spans a fine balance between the need to test for the acquisition of a clearly defined body of knowledge in the sciences and the evaluation of a range of less concrete but equally vital cognitive skills assessed within a humanities/social science environment. All these are evaluated within a framework that seeks to assess knowledge and understanding gained in the context of problem solving / reasoning / critical thinking ability. The science section of the test in particular is under constant tension from the competing interests of assessing specific subject knowledge gained and knowledge applied in a problem solving context. A further layer is added by the desire to integrate the three areas tested (chemistry 40%, biology 40% and physics 20%) where possible into hybrid units.

A further tension in both the test and the GEMP concept is the desire to attract more rounded applicants, including non-science graduates, while acknowledging that students must have command of the basic scientific concepts on entry in order for them to be successful in a four-year course.

Candidates are not recommended to any particular preparatory course of study or text book, but it is acknowledged that the majority will need to invest some time in preparing themselves for at least one section of the test. ACER produces a booklet of Sample Questions to enable candidates to make a realistic assessment of their strengths and weaknesses prior to undertaking the test. Expected standards are to first year university level in biology and chemistry with the necessary physics and mathematics to support this (ie approximately Year 12 level).

GAMSAT SCORING PROCESS

GAMSAT scoring is based on item response theory (IRT) methodology, a widely used procedure whereby items are calibrated to provide an estimate of the difficulty of each, expressed on an interval scale in terms of logits. Candidates' scores are expressed on the same interval scale as ability estimates which are then rescaled onto a GAMSAT scale in the range 0 - 100 for each of the three sections using the linear transformation:

$$\text{GAMSAT Section score} = (\text{estimate in logits} + 5) \times 10$$

The mean and standard deviation of Section I scores are used to standardise the scores on Sections II and III. The GAMSAT Overall score is calculated as a weighted average of the three section scores using the formula:

$$\text{Overall Score} = (1 \times \text{Section I} + 1 \times \text{Section II} + 2 \times \text{Section III}) / 4$$

The weighting of Section III is controversial and is seen by some to be a potential source of disadvantage to candidates who are non-science graduates. This could be overcome by reporting only Section scores in future. It is not possible to report separate scores for the physics, chemistry and biology components of Section III because of the

integrated nature of some of the units and also because numbers of items in each are not sufficiently great to allow meaningful scores to be derived.

Each Section II (Written Communication) task is marked by three independent raters, using a ten-point scale. Differences of 5 or more marks are considered discrepant and result in a fourth marking. The

three closest marks for each piece of writing are then summed. Interestingly, adjacent marks were given in 67% of cases in GAMSAT95, with a fourth mark being necessary in fewer than 10% of cases. Analysis of the relative difficulty of the topics (candidates selected from 5 prompts in each task) and the harshness of raters revealed that there was excellent consistency among raters and that difficulty levels were consistent across all topics. Perhaps the only surprise in the marking process was that the standard of writing was lower than expected for a graduate population.

PERFORMANCE OF GAMSAT95

Of the 815 candidates who sat GAMSAT95 males accounted for 50.3% and females 49.4%.

Table 1 indicates the performance of all candidates, including males and females separately, on GAMSAT95 and its components, in terms of mean and standard deviation of GAMSAT scores.

Table 1 Means and standard deviations for GAMSAT95 components (N = 815; Males = 410; Females = 403)

	All	Males	Females	
	mean	SD	mean	
	SD	mean	SD	
	mean	SD	mean	
	SD	mean	SD	
Total	58.35	5.59	59.25	8.57
Section I	58.37	4.58	58.37	5.58
Section II	58.37	4.58	58.17	3.58
Section III	58.37	4.60	56.17	6.56
Biology	57.18	9.59	59.18	8.55
Chemistry	57.39	5.58	58.81	0.35
Physics	58.59	8.60	49.95	6.69

Table 1 shows that males performed better than females on the test as a whole. This difference was largely accounted for by Section III, with a 3.7 point difference in mean scores. What was not expected was the fact that females did not outperform males in Section I and barely did so in Section II (0.3 points of difference). Also surprising was the fact that biology, the most verbal of the Section III components, showed over 4 points in favour of the males.

Table 2 Correlations Among Sections and Subsections

	Total	Un.Tot	Sec 1	Sec 2	Sec 3	Biol	Chem	Phys
Total	1.00	.97**	.75**	.54**	.84**	.72**	.68**	.45**
Un.Tot	.97**	1.00	.83**	.69**	.67**	.61**	.53**	.35**
Sec 1	.75**	.83**	1.00	.43**	.40**	.43**	.26**	.18**
Sec 2	.54**	.69**	.43**	1.00	.08*	.11**	.07-	.01
Sec 3	.84**	.67**	.40**	.08*	1.00	.80**	.85**	.59**

Biol.72**.61**.43**.11**.80**1.00.47**.28**
Chem.68**.53**.26**.07 .85** .471.00 .38**
Phys.45**.35**.18**-.01 .59** .28**.38**1.00

* - Signif. LE .05 ** - Signif. LE .01 (2-tailed)

The table shows Section I is not as highly correlated with Section II as might be expected. The somewhat low correlation of 0.43 suggests that the Written Communication is testing a range of skills and abilities not tested by the multiple-choice Section I items, which is a good result, since the validity of the combined Sections I and II must be greater than that for each one separately. The correlations between Section I and the three components of Section III are mixed. As might be expected, there is a reasonable correlation with biology, but fairly low correlations with chemistry and physics.

SOME DEMOGRAPHIC FEATURES OF GAMSAT95 POPULATION

Demographic features of GAMSAT95 should be viewed with some caution. It is acknowledged that the program, being a first in Australia, may take a few years to stabilise in terms of test taking population. It was obvious from the range of enquiries and registrations received for GAMSAT95 that there was a significant group of 'older' - (>30 years) people with a longstanding and frustrated desire to study medicine. The 27.2% of GAMSAT95 candidates in the over 30 age group may well not be repeated in future years, especially if it transpires that few of them are successful in gaining entry. Indeed it is fully expected that the eventual population will consist predominantly of candidates who have just finished or who are in the final year of study for their bachelors degree.

Age

The mean age of the candidates was approximately 25 years, with 61.5% aged 24 years or younger. 22% were aged 30 years or more, and 5.2% were older than 39 years.

The average performance by age showed little variation, except for a slight drop for the oldest age group (>39 years). Generally the performance of candidates in the younger age groups was more homogeneous than that of candidates in the older age groups. Younger candidates, who are still studying or who have recently completed formal studies, may have a decided advantage.

Average performance on Section I showed a slight but noticeable increase with age. This increase was even stronger for Section II. For Section III the average performance decreased significantly with age, and this pattern was observed in biology, chemistry and physics. This

would seem to support the interpretation that Sections I and II reflect maturity while for Section III the recency of the knowledge is of prime importance.

School Type

Distribution of candidates by school type was: Government 42.1%, Independent 31.8%, Catholic 21.5%, TAFE 1.8%, and Other 1.5%. (Missing was 1.3%)

In general, across the whole test and the three Sections, the average performance of candidates who attended independent schools was consistently better than that of candidates who attended other school types. Candidates from Catholic schools performed marginally better than did candidates from government schools. Candidates from Other and TAFE performed on average consistently worse than did candidates from the main school types.

Major subject area of first degree

Major subject studied during the first degree showed the strongest influence on test performance of any of the variables. The distribution of candidates by subject area of first degree was:

Biological Sciences 28.3%, Medical Sciences 22.9%, Allied Health Professions 22.4%, Arts /Social Sciences 13.2%, Physical Sciences / Mathematical Sciences 11.0%, and Other 2.2%.

These figures clearly show that the majority of candidates for GAMSAT95 (73.6%) were from discipline areas with strong links to medicine. It is anticipated that in future, as GAMSAT and the entry requirements for undertaking graduate medical study become more widely known, more candidates will be attracted from discipline areas less strongly linked to medicine. Obviously the challenge to the program is to increase the percentage of non-science candidates and to develop in a way that enables such candidates to do well on the test. One very pleasing result from GAMSAT95 was that the Arts/Social Sciences group featured in the top 200 scoring candidates in equal proportion to the numbers who sat the test.

Average performance on test Sections was in line with expectation. Candidates from discipline areas closely identified with a GAMSAT component did better on that component than did candidates from the other discipline areas. Thus the Arts / Soc. Science candidates did best on the verbal components, Sections I and II, whilst candidates with a Physical and Mathematical Sciences background did better in chemistry and physics. Candidates from the Biological and Medical Sciences did better than other candidates on the biology component.

Highest Degree Level

The majority of candidates (74.8%) had only a first degree or were still studying for a first degree. 17.5% of the candidates had an Honours degree and 6.9% had a higher degree (Masters or PhD). (Missing was 0.7%)

A surprising observation was that honours graduates performed

consistently better than both bachelors and higher degree graduates. Candidates whose highest qualification was a bachelors degree performed on average slightly worse than candidates with higher degrees.

CONCLUSION

GAMSAT represents an exciting move for ACER into the assessment of higher level cognitive and curriculum skills and abilities and provides a new opportunity for the development of staff expertise in testing at the tertiary level.

BIBLIOGRAPHY

Morgan, G., Congdon, P. & McCurry, D.
Report on GAMSAT95. Camberwell, Vic. ACER, June 1995

GAMSAT Information Booklet 1995-96
Camberwell, Vic. ACER

Graduate Medicine
(booklet for intending students of the Graduate-Entry Medical Course)
The Medical School, Flinders University of South Australia, Adelaide

ACKNOWLEDGMENTS

Dr Geoff N Masters, Associate Director (Measurement), ACER, Melbourne

Professor Ann Sefton, Associate Dean, Curriculum Development, Faculty

of Medicine, University of Sydney

Dr Jillian Teubner, Medical Education Unit, Flinders Medical Centre,
Adelaide