

THINKING SKILLS TESTS FOR UNIVERSITY ADMISSION

Robert Harding

Thinking Skills Tests for University Admission

Robert Harding

University of Cambridge Local Examinations Syndicate (UCLES),
Interactive Technologies in Assessment and Learning Unit (ITAL),
1 Hills Road, Cambridge CB1 2EU, United Kingdom
email: r.d.harding@ucles-red.cam.ac.uk

Background

The relatively large number of university applicants predicted multiple A-grades in end of school A-level examinations when they apply has led to growing interest by several Universities in special additional tests for admissions purposes. The UK Government is also concerned about access to university courses, and in May 2003 Secretary of State for Education and Skills Charles Clarke requested Professor Steven Schwartz, Vice-Chancellor of Brunel University, to conduct an independent review¹. The enquiry's concerns about widening access to higher education has also led to a search for additional admissions indicators. Tests based on Thinking Skills appear promising for both purposes, and CAA has a major contribution to make in delivering them successfully.

UCLES has taken an interest in Thinking Skills tests since the late 1980's/early 1990's. Tests such as MENO (named after a disciple of Plato; the name is not an acronym) were developed. Although MENO did not continue, the OCR² Board's Advanced Subsidiary (AS) General Certificate of Education qualification in Critical Thinking and CIE's² AS in Thinking Skills are derived from it. Two examples of the literature that is available to support preparation for such tests are Fisher (2001) and Thomson (1996).

The need for special admissions tests is not new or unique to the UK, and there are a number that are well established. Examples are:

- The United States Medical Licensing Examination (USMLE)
<http://www.nbme.org/about/about.asp>

¹ <http://www.admissions-review.org.uk/>

² Both the Oxford Cambridge and RSA (OCR) Board and Cambridge International Examinations (CIE) are part of the UCLES Group.

- The Law School Admission Test (LSAT)
<http://www.lsac.org/>
- Graduate Medical School Admissions Test (GAMSAT)
<http://www.acer.edu.au/tests/university/gamsatuk/intro.html>

These are designed for entry to professionally oriented courses and are often quite expensive.

In the USA, where there is no national system of examinations at the end of secondary education, the College Board's SAT is widely used to provide a common standard when colleges are faced with making decisions on the basis of GPA's (Grade Point Averages) which are awarded locally and which are not comparable across a huge nation. When originally created in the 1920's, the SAT was intended to be a test of aptitude and is sometimes mistakenly thought to be so still, but this is no longer the case:

"The SAT is a three-hour test that measures verbal and mathematical reasoning skills students have developed over time and skills they need to be successful academically."

<http://www.collegeboard.com/student/testing/sat/about/SATI.html>

There is also a SAT II test, which provides more subject/curriculum based measures of achievement. The SAT's are not the only such instruments in the USA. For instance ACT, based in Iowa City, is a significant competitor via the ACT Assessment; it is curriculum based and covers English, Maths, Reading and Science, and has never claimed to measure aptitude. (<http://www.act.org>). Schneider and Dorans (1999) report a correlation between SAT and ACT of around $r = 0.9$.

The purpose of such tests is to predict achievement in further studies. Geiser and Studley (University of California, 2001) looked at ability of SAT I and II to predict college achievement (as measured by freshman Grade Point Average). Considering scores on either of the SAT tests in addition to high school grades added to the amount of variance in freshman Grade Point Average that could be explained. However, SAT II score was found to be the best single predictor, then high school Grade Point Average and then SAT I score. It has previously been argued that the SAT I is likely to be more helpful than the SAT II in identifying students with potential in schools that perform less well because the SAT I is more curriculum independent. The University of California analysis did not support this and found SAT II results to still be a better predictor of freshman grade than the SAT I even when analysed by school background.

The SAT I has also been tested in the UK (McDonald et al, 2001).

This paper is not however about these tests and their properties but rather about how the construction and administration of such tests interacts with the technology opportunities that CAA brings. This paper describes UCLES' experiences in these aspects with the two tests with which we have become involved, the BioMedical Admissions Test (BMAT) and the Thinking Skills Assessment (TSA).

Development of the BMAT

In 1999 the University's Medical and Veterinary Schools asked UCLES to assist in developing a special test for applicants to assist admissions staff in selecting from amongst the large number of highly qualified candidates who apply to Cambridge. Our Research and Evaluation Division (RED) constructed the MVAT (Medical and Veterinary Admissions Test). It consisted of 3 components: two multiple choice papers lasting 30 minutes each, one on thinking skills, the other on basic scientific and numeracy skills, and a free response short essay paper lasting one hour. After a small-scale trial in 1999, the MVAT was run for all Cambridge Medical and Veterinary applicants in November 2000, 2001 and 2002 with candidature increasing from about 1600 to 1800. Other Universities took a strong interest and asked to use the test. For the November 2003 test, Oxford and University College London (UCL) Medical Schools joined the pilot, and the test was run with a modified format and had its name changed to BMAT (BioMedical Admissions Test). The time for the Thinking Skills component was increased to 1 hour and for the essay component reduced to 30 minutes. The BMAT was administered to about 4000 applicants at their schools or other centres in November, in advance of any admissions interviews.

The test itself appears to be performing very well and provides useful information: RED is working on a detailed evaluation, and it is hoped to present highlights of the findings at CAA 2004. A website is available giving details of the test with example and practice materials³. The test itself is delivered entirely on paper at present; it has been described in order to give the context in which the work described is being done, and this paper will not give any further details.

Development of the TSA

As for MVAT/BMAT, the demand for a "Thinking Skills Assessment" to assist in making admissions decisions originated from Cambridge University admissions staff. An item bank of former Thinking Skills questions (items) was built up for this purpose. Unlike the BMAT, the TSA was seen as a test that would form part of the admissions interview process, so it was to be taken by applicants during their interview visits to Cambridge. This has the advantage in the Cambridge context of allowing the use of the test to be made on a college-by-college and subject-by-subject basis. The test consists of 50 multiple-choice questions and takes 90 minutes. In December 2001 about 289 Computer Science applicants took the test. This expanded to about 472 in December 2002 with more Colleges and more subjects taking part.

Up to this point the objective of the TSA work was the development and evaluation of the test itself, but in January 2003 UCLES decided to add a second objective, that of experimental online delivery of the test. This

³ <http://www.bmat.org.uk/>

software was developed specifically for UCLES as a prototyping system, to our specifications.

Both objectives were successfully achieved: there was greatly enhanced take-up of the test, 23 Colleges taking part involving 4 main subjects (Computer Science, Engineering, Natural Sciences and Economics), and the administration procedures were based around the online system we had developed. Altogether a total of 1551 tests were administered, 1114 paper tests and 437 online tests. An especially valuable feature was the administrative website, which was used for making entries (registration), and returning results. On-line tests were marked automatically, and paper tests were marked using scanning technology with intelligent character recognition. A website is available giving details of the test with example and practice materials⁴.

The software development work included the development of an item banking system capable of delivering master copies of tests whether in on-line or paper format, and new items are now being commissioned using this system. Authors can key in their items by remote access.

Outcomes

The section refers specifically to the TSA.

A “customer satisfaction” survey has been carried out amongst the Colleges using the test, and the feedback is largely positive. 26 Colleges were using the TSA. 19 of these answered the survey, and 16 agreed, some of them strongly, that the administration of the test ran smoothly. Since web based communications and email were used extensively in the administration, this is positive news for CAA. 16 Colleges said they found that receiving the results online was useful.

Interestingly though, 7 Colleges amongst those taking the test on paper were not convinced by the attraction of CAA and have said that they will be staying with the paper version; only 3 said they were going to use the online version in future. On the other hand, 2 users of the online tests said they were going to use the paper version in future. Perhaps the CAA community will draw a small measure of encouragement that at least the net flow is in the direction of CAA, though it hardly amounts to evidence that the community of users is clamouring for online testing. The reason for this we think is that some of their networks are not sufficiently reliable for admissions staff to be sufficiently sure that it can support a high stakes test, and others find the provision of a sufficiently large and secure computer room problematic.

For the usefulness of the test itself in selecting applicants, we have to turn to earlier admissions rounds because it is only at the end of their first year at

⁴ <http://tsa.ucles.org.uk/index.html>

University that we can obtain any objective measure of a student's success. We have done an analysis of the June 2003 Computer Science examination marks of those approximately 50 successful applicants who sat the TSA in December 2001, correlating these with their TSA scores. The results show a better correlation than for A-level or interview scores. Evaluation will continue both of the TSA and the BMAT, and it is hoped to present highlights of the findings at CAA 2004.

Comparability between paper and online testing

One of the key issues for CAA is the comparability of online and paper versions of tests. As McGuire and Youngson (2002) point out, with current technologies and if automatic marking is wanted, it is often not possible to set a question that is the same task both online and on paper. For example, a question in mathematics starting "Prove that ..." cannot be marked automatically. McGuire and Youngson found however that when identically worded questions were considered, and when presentation was comparable (e.g. one question per screen, one question per page), there was no significant difference between these two media as far as test difficulty was concerned.

Our results for the December 2003 administration of the TSA gave similar findings. All TSA questions are multiple-choice, and have been consciously designed from some years back to be suitable for both online and paper media.

Table 1 shows the mean and standard deviation of scores for all candidates who took a test on paper, and separately for all candidates who took a test online. The figures are very similar and suggest that mode of delivery generally had little effect on scores.

Table 1: Means and standard deviations by mode of delivery (paper or online) of test

Delivery Mode	Score type	N	Mean	Std Dev
Paper	Total	1114	60.4	8.3
	Prob. Solv.	1114	62.5	10.2
	Crit. Think.	1114	59.0	10.1
Online	Total	437	61.0	8.1
	Prob. Solv.	437	63.4	10.2
	Crit. Think.	437	59.5	9.6

The future for special admissions tests

UCLES “Assessment Directorate”, of which the ITAL Unit forms a part, has as one of its main objectives to provide a development ground for innovative assessment projects. It is appropriate to include admissions testing within the Assessment Directorate, as the administration of these tests does not fit the familiar operational models of A-levels. There is also considerable overlap between the operations of the BMAT and TSA. We have therefore formed the “Thinking Skills for Admissions Testing” (TSAT) project.

The overall strategy for the project is based on three main areas of activity: test production, administration operations, and test development, evaluation and research. All of these will be essential if this new area of activity is to be successful. Of particular interest to the CAA community will be the way that websites are used to disseminate information about the tests, provide practice, preparation and other study materials, an administrative interface, and not least, on-line versions of the tests themselves. CAA offers many advantages in running this kind of “additional information” test, and in the short term there is a great deal of interest.

Longer term the prospects are less certain. If for example appropriately scaled A-level (or Tomlinson diploma) marks are released to admissions officers instead of grades (as at present), perhaps the demand for special tests will fall off. Against this, BMAT for example is designed to test rapid logical thinking and decision-making as these are desirable qualities for future medics, qualities that are not tested by a conventional A-level. Similar requirements may exist in other subjects, e.g. Law. TSAT type tests seem in keeping with the thinking that has come from the Schwartz consultation in that they provide an alternative way for applicants to assess themselves for intellectual qualities, and the skills they assess, while essential for high level study, are not only developed through school based academic study.

References

Assessing ICT Assessment in Mathematics, G R McGuire, M A Youngson, <http://itsn.mathstore.ac.uk/articles/maths-caa-series/index.html>, , , (2002)

Critical Thinking: An Introduction, Fisher, Alec (Edition 2001) ISBN: 0521009847 Publisher: Cambridge University Press, The Edinburgh Building, Shaftesbury Road, Cambridge, CB2 2RU, United Kingdom <http://uk.cambridge.org>

Critical Reasoning: A Practical Introduction, Thomson, Anne (Edition 1996) ISBN: 0415132053 Publisher: Routledge, Taylor and Francis Books, 11 New Fetter Lane, London, EC4P 4EE, United Kingdom <http://www.tandf.co.uk>

Differential item functioning for minority examinees on the SAT, Schneider, A P and Dorans, N J (1990). *Journal of Educational Measurement*, **27**, 1, 67 – 81.

UC and the SAT: Predictive Validity and Differential Impact of the SAT I and SAT II at the University of California, Geiser, S and Studley, R (2001)
California, UC Office of the President.

A Pilot of Aptitude Testing for University Entrance, Slough, NFER.McDonald, A S, Newton, P E and Whetton, C (2001)

The work reported in this paper is the joint work of a team in UCLES, and the author acknowledges the many contributions made by colleagues to this paper. In particular, thanks to Nick Raikes for the statistical work, to Vicki Crisp for assistance with background research, and to Maura Bourdon and Rosemary Whittle for their critical reading of drafts and suggestions for improvements.