Published in Assessment Matters

Does Psychometric Testing offer a solution to university selection?

Author: Angus S. McDonald

Introduction

As the number of young people applying for higher education places has increased dramatically over the last decade, so the methods by which universities select applicants has come under closer scrutiny. At present, A-levels remain the most common route into higher education. However, as the proportion of students attaining 'A' and 'B' grades has increased, A-levels have become less useful to admissions tutors as a means of discriminating between students. In the absence of clear guidance from A-level grades, it is not always clear how admissions tutors make the difficult decision of which students to offer places and which to reject.

Statistics published by the Sutton Trust in 2000 showed that students from independent schools were over-represented in the highest-ranked universities, even after allowing for differences in A-level grades. This report suggested two main contributory factors that may have led to this situation: a lower number of students from less affluent (typically state school) backgrounds applying to leading universities, and inadequacies in the admissions system.

From the discussion on university access which followed, the SAT was put forward as one way of identifying a student's potential for higher education. Supporters of the SAT drew on claims about its ability to predict success in higher education irrespective of a student's educational and social circumstances. Against the background of controversy surrounding the British admissions system, the SAT naturally appeared appealing.

Due to the lack of recent research on the SAT in Britain, the Sutton Trust commissioned the NFER to conduct the research summarised here.

What is the Schalastic Assessment Test?

The College Board was founded in America in 1900, with the goal of reconciling the differences between colleges' admissions procedures. The tests offered by The College Board from 1901 were open-ended assessments. It was not until 1926 that the test recognised as the SAT was first offered, its development having been influenced by the extensive use of objective testing for the selection of army recruits in the First World War. In 1947 The College Board helped found Educational Testing Services (ETS) who have since been responsible for the development of the SAT.

Apart from minor modifications, the SAT remained largely unchanged until 1994 when it underwent a major revision. Currently, the SAT assesses verbal and math reasoning ability. The verbal SAT consists of:

- analogies measuring reasoning skills and knowledge of vocabulary (19 questions);
- sentence completions measuring logical relationships and vocabulary in the context of the sentence (19 questions);
- four reading passages of between 400 and 800 words each, with questions assessing reasoning skills, literal comprehension and vocabulary in context.

The math SAT consists of 35 multiple-choice questions, 15 quantitative comparison questions and 10 open-ended or 'student-produced' responses, which cover he areas of arithmetic, algebraic and geometric reasoning.

Students are allowed 1 hour and 15 minutes to complete each of these sections.

Aims and methodology

Although appealing as an idea, two fundamental questions needed to be answered if the SAT was to be of potential use. The first of these was whether the SAT measured factors distinct from those assessed by A-levels – if not there would be little purpose in studying it further. The second concerned the 'fairness' of the SAT – whilst many of its proponents claim SAT performance to be relatively independent of social and educational experiences, this has been hotly debated in America for many years. The pilot study conducted by the NFER set out to provide evidence on these two points, although as it was an initial study it did not address whether the SAT predicted long-term success in higher education and beyond.

During the spring and summer of 2000, almost 1300 A-level students took a short version of the SAT I: Reasoning Test. Students were sampled from low- and high-achieving schools on the basis of their GCSE performance, with the latter sample also including students from selective independent schools. The SAT contained verbal and math sections, with students being allowed 30 and 40 minutes respectively to complete these. Prior to taking the test, students were given a brief coaching session offering tips for taking the SAT and practice questions. A-level results were collected when published.

A literature review was also conducted to look at broader issues surrounding aptitude testing for higher education and to report on a number of current initiatives in British universities.

Findings

The first finding to emerge from the study was that the SAT and A-levels were assessing relatively distinct constructs, with there being around 25 per cent of shared variance between the two. This figure is remarkably similar to that obtained from previous work on aptitude testing for university entrance in Britain, conducted in the 1970s. Interestingly, GCSE scores correlated slightly higher with SAT scores than A-levels did; the two sharing around 30 per cent of variance. Despite the closer proximity between students taking A-levels and the SAT, this association

may be due to GCSEs reflecting a wider overall course of study, as students typically take English, maths and at least one science subject at GCSE.

The next question concerned whether, as has been claimed by some of its supporters, SAT scores were minimally affected by social and educational factors. The results did not support this claim. Firstly, SAT scores varied by background factors in much the same way as A-levels did. For example, students from independent schools achieved the highest A-level and SAT scores, followed by the high-achieving and then the low-achieving schools. Similar correspondence between A-levels and SAT scores was seen for sex, ethnicity, parental socio-economic status and students' intended destination after A-levels. Secondly, an analysis using multilevel modelling found that the association between SAT scores and A-levels remained relatively constant, regardless of background factors. There was, therefore, no evidence that the math and verbal sections of the SAT were measuring constructs more independent of students' social and educational experiences than A-levels.

The modest correlation between SAT scores and A-levels meant that, when individual students were considered, some attained high SAT scores but only modest A-level grades. This was most noticeable in the low-achieving sample, where around five per cent of students achieved SAT scores sufficiently high to be considered for an American Ivy League college, although only one achieved three As at A-level – the effective entry requirement for Oxbridge. Conversely, some students had high A-level grades but lower SAT scores. If A-levels and SAT scores were combined, a greater number of students could have been considered for university places. There was some evidence that this may have favoured students from the lower-attaining schools.
Further analyses showed that the SAT was a reliable assessment of reasoning in a sample of British students and that it showed minimal item-level bias according to sex or sample (high- and low-achieving and independent schools). Further, British students scored comparably to their American counterparts on the SAT, despite having minimal preparation and being less familiar with the multiple-choice format.

Implications
So what are the implications of this work for access to higher education in Britain?

The finding that the SAT is measuring a construct distinct from that assessed by A-levels suggests it is worth further study. There is now a need for longer-term work relating SAT scores to university performance. Only through rigorous predictive studies will it be possible to tell whether aptitude testing could have a role to play in access to universities, particularly whether high SAT scorers from lower-attaining schools go on to perform well.
Less promising are the results showing that SAT scores are no more independent of social and educational factors than A-levels. Although this work was based on limited data, it is in accordance with evidence from America where the use of the SAT has been challenged legally in some states because of its apparent bias.

It emerged during the course of this work that some universities are already looking at additional assessments that may enable them to widen access. For example, initiatives at Oxford and King's College London are developing assessments that allow students to demonstrate their potential for learning. Another university is focusing more on identifying personality characteristics and other 'softer' predictors of success, attempting to provide a more objective assessment of the characteristics evaluated through interviews. As all this work is in its early stages, it remains to be seen how effective it will be.

More fundamentally, research such as this forces us to re-examine the purpose of our higher education system. If its purpose is purely to produce academically able students then SAT-like tests, which conform to a very traditional view of intelligence, may have much to offer. An alternative perspective, and one that appears increasingly significant now that around one-third of all young people go on to higher education, takes a broader approach to the skills people need to be successful in the modern world and what they can offer in return.

The literature review found that use of the SAT by American colleges has changed over time, with its value being increasingly questioned. Where it is used admissions tutors often 'contextualise' the scores, accepting that SAT scores are not absolutes but need to be considered in relation to a student's social and educational background. Here, the goal is not to produce straight-A students, but to look at the potential a college education may give young people to contribute to wider society in the future. These colleges take what could be called a 'mature' approach to SAT scores, developed from many years of experience and by taking a broader view of the role of education in society. If a version of the SAT is ever adopted in Britain it is clear that experiences in America have many valuable lessons to teach us.

Bibliography
McDonald, A.S., Newton, P.E., Whetton, C. and Benefield, P. (2001) Aptitude Testing for University Entrance: A Literature Review. NFER: Slough
McDonald, A.S., Newton, P.E., Whetton, C. and Higgs, S. (2001) A Pilot of Aptitude Testing for University Entrance. NFER: Slough
This work was commissioned by the Sutton Trust and conducted by the NFER. Copies of the reports may be obtained from the NFER Library: +44(0) 1753 574123, www.nfer.ac.uk. Angus S. McDonald is now a Consultant Psychologist with Team Focus Limited: +44(0) 1628 637338, www.teamfocus.co.uk.