Ability, attainment and affluence – relationships between the SAT®, A levels and student characteristics

Catherine Kirkup, Rebecca Wheater, Ian Schagen and Jo Morrison
National Foundation for Educational Research

Paper presented at the IAEA Annual Conference, Cambridge, 8-12 September 2008

Abstract

In a longitudinal study to examine the validity of the SAT® for use in admissions to higher education in the UK, relationships between SAT® score, UK national examination results ("A levels") and background characteristics were analysed. Using an affluence measure derived from students' survey responses, it was found that SAT® scores on two components (Critical Reading and Writing) tended to be higher for affluent students compared to less affluent students with similar A level and GCSE results. The level of entry points of students' HE courses were related to factors at both the individual student level and the school level. The former were primarily their attainment at A level and GCSE, but also factors such as ethnicity and affluence. The SAT® may offer some incremental information to aid the selection of HE candidates over and above that provided by performance at both GCSE and A level, and there is some evidence that SAT® score may be a more important predictor for students in schools which do less well at GCSE.

1 Introduction

In September 2005 the National Foundation for Educational Research (NFER) began a five-year research study to examine the validity of the SAT Reasoning Test™ in higher education admissions[1]. This paper reports on one aspect of the interim analyses - the relationships between SAT scores, attainment in A level examinations and background characteristics of the students related to levels of affluence. For more detailed accounts of the first stages of this project see Kirkup *et al*. (2007 and 2008) and Whetton *et al.* (2007).

2 Background

Higher education brings considerable benefits to graduates in terms of salary, job security, employment opportunity, and so on, although these will vary somewhat according to the course studied and the institution attended. The government is committed to achieving social inclusion within higher education. However, although the number of students entering higher education in the UK has grown enormously in recent years, some groups are still under-represented. Over the period 1994-2000 young people living in the most advantaged 20 per cent of areas were between five and six times more likely to go into higher education than those from the least advantaged 20 per cent of areas (HEFCE, 2005). With an increasingly large number of highly-qualified candidates, some universities have introduced additional

[1] The study is co-funded by the NFER, the Department for Innovation, Universities and Skills (previously the Department for Education and Skills), the Sutton Trust and the College Board.

admissions tests to select candidates for courses where there is heavy competition for places.

Although 'prior educational attainment remains the best single indicator of success at undergraduate level' (DfES, 2004) it is recognised that for some students, their true potential may not be reflected in their examination results due to social or educational disadvantages. The ability of A level[2] grades to predict degree outcomes has been demonstrated using a large data set (Bekhradnia and Thompson, 2002). Whilst some evidence regarding the validity of admissions tests within the UK context has been published (Bell, 2005, Emery, 2007a, 2007b), to date such research has generally only been possible with relatively small numbers of students. Amongst its wider recommendations the Admissions to Higher Education Steering Group (DfES, 2004) encouraged the commissioning of research to evaluate the ability of aptitude tests to assess the potential for higher education.

The principal previous study underpinning this current research is the pilot comparison of A levels with SAT® scores conducted by NFER for The Sutton Trust in 2000 (McDonald *et al.*, 2001a, Whetton, 2001). The study found that the SAT® was only modestly associated with A level grades, which suggested that it was assessing a distinct construct from A levels. However, there was no evidence that the association differed according to background factors such as ethnicity, parental socio-economic status or overall achievement of the school. Although the SAT® has been relabelled as a 'reasoning' test, it is still generally perceived as a test of academic aptitude. Implicit within the term 'aptitude' is the concept of the predictive facet of validity, i.e. to measure an individual's potential for obtaining a certain goal, in this case, the successful completion of a university course. For a detailed discussion of aptitude testing for university entrance see also the literature review conducted by McDonald *et al.* for the Sutton Trust (2001b).

The primary aim of the study is to examine whether the SAT® and A levels together are better able to predict university participation and outcomes (degree success) than A level results alone. Another issue to be addressed is whether the SAT® can identify students with the potential to benefit from higher education whose ability is not adequately reflected in their A level results (because of their economically or educationally disadvantaged circumstances). Until degree outcomes for students in the sample become available in 2010, it will not be possible to answer the main research question. In the early phases of the research the analysis is therefore focussed on the relationships between SAT® scores, A level scores/grades, prior attainment at age 16 and background characteristics of the student sample.

Score differences on the SAT® between different ethnic and different socio-economic groups are well documented but such differences are also found in most other measures of educational achievement (Camara and Schmidt, 1991). An important

[2] A Level - short for Advanced Level. In England A Levels are studied between the ages of 16-18 years. The first year of A Level is called AS Level and the second year is called A2. Generally students take examinations at the end of both years. The curriculum is set by the government but the examinations and specifications (syllabuses) are set by individual boards. The examination is subject based with a separate grade for each curriculum area. Typically students study only three or four subjects rather than the broad curriculum typical of most other countries. All subjects are graded from A-E and typically students need 3 or 4 good A level grades to gain entry to a top UK university.

aspect of the research is whether the predictive power of the SAT® differs across different sub-groups of students and to consider the utility of such a test if such differences occur. In the United States, although high school grades are often seen as the slightly better predictor of college grades, it is reported that the SAT® adds to their predictive power to a statistically significant degree, and may be a more accurate predictor for some groups of students with discrepancies between grades and SAT® scores (Kobrin *et al*., 2002). In another study published by The College Board (Stricker et al, 2002) it is suggested that, although educational disadvantage is related to ethnicity and socioeconomic status, it may be possible and helpful to distinguish educational disadvantage as a separate construct. Unfortunately, subgroup performance differences (other than gender) are difficult to examine in detail within the current research because of the relatively small numbers of students in some sub-groups.

3 Sample and methodology

In autumn 2005 approximately 9000 students in English schools and Further Education colleges took the SAT Reasoning Test™ during the final year of their two-year A level courses. (For most students this is the academic year in which their 18th birthday occurs.) The SAT® comprises three main components: Critical Reading, Mathematics and Writing; for a full description of the test and the scoring metric see www.collegeboard.com.

In January 2007 the SAT® data for students in the sample was matched with the 2005/06 National Pupil Database supplied by the Department for Education and Skills (DfES). The dataset included their A level results, prior attainment at age 16 (GCSE[3] results) and, for students educated within the maintained sector, background characteristics from the Pupil Level Annual School Census (PLASC) data. The number of students with valid data on all three main variables (SAT® scores, A levels and GCSEs) was 8041. A comparable 'national population' was derived from the same National Pupil Dataset by extracting those students taking two or more A levels. Additional background characteristics were also obtained from responses to two optional questionnaire surveys carried out in spring and autumn 2006.

Key background characteristics of the main sample are shown in Table 3.1. These details were obtained by combining information from the PLASC data for students from maintained schools with information supplied by individual FE colleges and independent schools.

[3] GCSE (short for General Certificate of Secondary Education). GCSEs are the main examinations taken by UK students aged 16 at the end of statutory secondary education. These are subject based, with students studying eight or nine subjects typically.

TABLE 3.1: Background characteristics of the main sample

		Main sample		National population*	
		N	Valid per cent	N	Valid per cent
Sex	Male	3692	45.9	98625	45.6
	Female	4349	54.1	117718	54.4
Ethnicity	Asian or Asian British	670	9.1	7799	6.9
	Black or Black British	117	1.6	2243	2.0
	Chinese	116	1.6	996	0.9
	Mixed	145	2.0	1392	1.2
	White	6212	84.4	93732	83.2
	Other	104	1.4	6499	5.8
	Missing	677	-	103682	-
Free school meals eligibility	No	5953	96.1	114058	97.2
	Yes	243	3.9	3250	2.8
	Missing	1845	-	99035	-
Type of institution	Comprehensive	4200	52.2	99280	45.9
	Grammar	1701	21.2	19790	9.1
	Independent	1800	22.4	32544	15.0
	FE college	340	4.2	64729	29.9
Total		8041	100	216343	100

** Candidates entered for 2+ GCE A levels in 2005/06 (source: DfES)*
Valid percentages exclude missing data. Due to rounding, percentages may not sum to 100.

In the initial analyses the main study variables were A level scores, GCSE scores and SAT® scores. In addition to simple descriptive statistics of sub-group performance and correlations between the main study variables, various regression models were employed to explore more complex relationships. For example to explore what factors predicted higher or lower than expected performance on the SAT, once overall A level performance and prior attainment were taken into account.

4 Results

In the analyses that follow, the attainment data for students in the sample was taken from a dataset supplied to the NFER by the DfES. Further details of the scoring systems for both A level and GCSE qualifications and information about the discounting process (used to avoid double counting of qualifications such as A and AS levels) can be found on the DfES website (DfES, 2006).

4.1 Descriptive Statistics

Table 4.1 shows the sample and national means for the key attainment measures. The score distribution is slightly skewed towards the upper range compared to the national population of A level entrants taking two or more A levels, probably because of the

high number of students in the sample from grammar schools and independent schools. However, the distribution of scores broadly covers the same range as the population and therefore the sample contains sufficient cases from all areas of the population to enable reasonable conclusions to be drawn.

TABLE 4.1: Mean attainment scores – main sample

	Main sample		National population	
	Mean	**s.d.**	**Mean**	**s.d.**
Total A level point score	**848.6**	260.4	**808.4**	235.8
Total GCSE point score	**489.9**	80.1	**469.0**	107.6
Average GCSE point score	**47.4**	6.0	**46.4**	5.5
	n = 8041		max n = 216343	

All values significantly different at the 5 per cent level.

The SAT® means achieved in the sample were roughly comparable with US means and the individual items functioned in a similar way for students in England and the US. Breakdowns of the main study variables by background characteristics found sub-group differences by gender, by eligibility for free school meals (FSM), by ethnicity, by English as an additional language (EAL), by socio-economic groups and by type of school/college attended. (For full details see Kirkup *et al* 2007.)

4.2 Exploring the relationships between the main study variables

Correlations between GCSE scores (total and average) and A level total score and between GCSE and A level scores and SAT® scores were calculated. The highest correlation with average SAT® score (across Critical Reading, Maths and Writing) was average GCSE score (0.70), followed by A level total score (0.63) and GCSE total score (0.54).The correlation of total A level points with average GCSE score (0.76) was higher than with the total GCSE score (0.58). It is likely that this is because the number of GCSEs entered can vary widely and does not always reflect the ability of the student whereas at A level there is far less variation in the number of A levels attempted. The high correlations between SAT® scores and attainment at GCSE and A levels are not unexpected given that each of these is measuring educational ability, albeit different aspects and in different ways. The relationship between A levels and SAT® scores is complicated in that each of these measures is associated with prior attainment at GCSE. Controlling for average attainment at GCSE, the partial correlation between SAT® and A levels was 0.23. This suggests that the underlying constructs that are being measured are somewhat different. This may indicate a potential for the SAT® to add to the prediction of higher education outcomes from A levels, although the increment is likely to be relatively small. This issue will be examined further when degree outcomes are available for students in the sample.

An issue of particular interest was whether particular types of student performed better in the SAT® than would be predicted from their A level and GCSE results. A regression model was run, with average SAT® score as the outcome, controlling for both A level total points and average GCSE points. Similar regressions were run for the individual components of the SAT®. The analysis showed that some groups of

students appeared to perform less well on the SAT® than expected: females compared to males and Asian, Chinese and those whose ethnicity was unknown compared to White students; whereas students in grammar schools did better than might be expected compared to comprehensive students.

One of the research questions is to explore the ability of the SAT® to identify students with the potential to benefit from higher education whose ability is not adequately reflected in their A level results because of disadvantaged circumstances. Therefore a particular variable of interest was the free school meals indicator (FSM), which is often viewed as a proxy for disadvantaged circumstances. This indicator was not a significant variable when attempting to predict average SAT® performance, i.e. there was no tendency for students eligible for free school meals to perform better (or worse) in the SAT® than in the other two attainment measures. However, when a regression model was run for each of the three main SAT® components separately, FSM students did better than would be expected on the SAT® Critical Reading component compared to students that were not eligible (i.e. they tended to achieve higher scores than non-eligible students with similar GCSE and A level attainment.) In the regression analysis model to predict higher or lower performance in the SAT® Maths component, FSM students did worse than expected.

However, for this dataset, FSM is likely to be a poor indicator of social and economic disadvantage. Only a small proportion of students in this relatively high-performing group were actually eligible for FSM, and within the non-FSM category there is likely to be wide variation in affluence / disadvantage.

4.3 Improved measures of affluence / disadvantage

In order to explore students' SAT® outcomes according to more subtle measures of disadvantage (i.e. superior to the simple FSM indicator), additional items of background data were used, where available, to supplement the broad background characteristics already examined. These additional items were derived from a number of different sources as outlined below. With this extra information it was felt it might be possible to understand better the complex relationships between students' backgrounds, the results they achieve and their success at gaining entry to higher education.

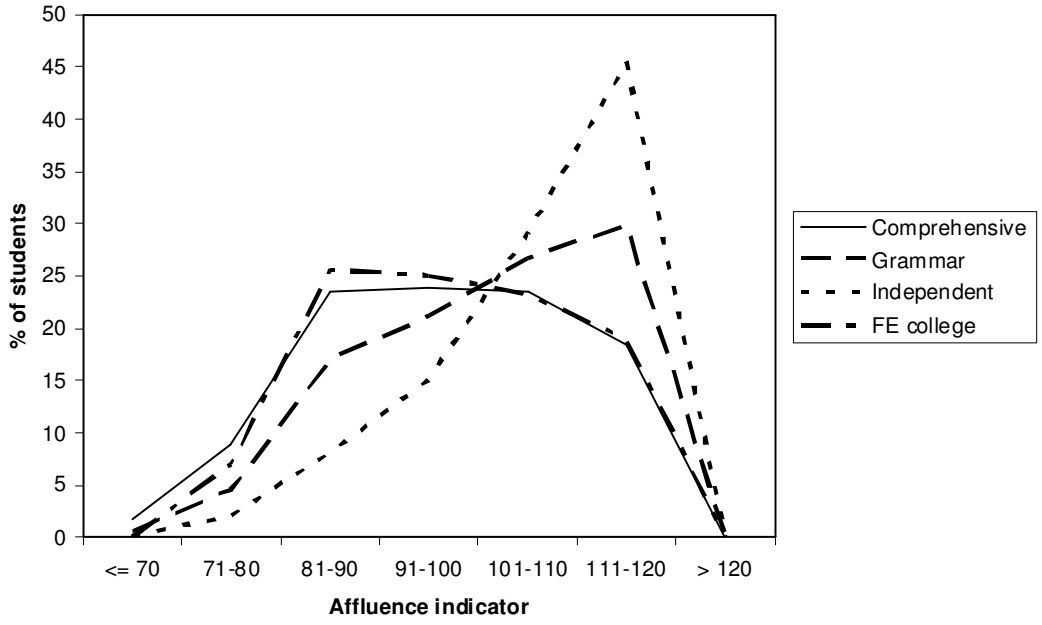
School level measures: the school or college GCSE band[4]; the school or college A level band[5].

IDACI (Income Deprivation Affecting Children Index) This index measures the proportion of children under the age of 16 in an area, who are living in low income households. The IDACI indicator ranges from 0.00 to 1.00 with 0.14 being around average nationally (see DCSF, 2005). Higher scores indicate greater deprivation. In our sample the average for the known sample (5858 students) was 0.13. The IDACI of those students eligible for free school meals (232 students) was 0.33.

[4] School GCSE band: students' total GCSE point score is averaged across the school, and schools are divided into 5 bands, of 20% each, ranging from the lowest attaining schools to the highest. This relates to the school where the student took their GCSE exams.
[5] School A level band: students' total A level point score is averaged across the school, and schools are divided into 5 bands, of 20% each, ranging from the lowest attaining schools to the highest.

Affluence indicator Student questionnaire responses were included in a factor analysis to produce a measure of affluence. The items used to derive this indicator were: socio-economic group (based on the occupation of highest-earning parent), home owned or rented, number of siblings, books in the home and level of education of both mother and father (see Appendix 1, Kirkup *et al*., 2007). The Cronbach's alpha[6], a measure of reliability, of the affluence indicator was 0.54, which is respectable for a scale derived from six questionnaire items. With a mean indicator of 100 for the questionnaire sample as a whole (5059 cases), students from comprehensive schools had the lowest affluence indicator (96.8) and students from independent schools had the highest (106.3), see Figure 1.

Figure 1: Affluence indicator by institution type



4.4 Analysis of SAT® scores including affluence / disadvantage measures

The analysis of SAT® scores controlling for pupils' GCSE and A level results was rerun in two regression models: the first using IDACI as well as school-level indicators (n = 5815); and the second with the affluence indicator plus school-level indicators (n = 4806). The results from the analysis are presented in Tables 4.2 and 4.3 respectively.

For categorical variables, presented below the dashed line, the change in SAT® score is the difference between categories, e.g. boys compared to girls. For the non-categorical attainment variables, above the dashed line, the change in SAT® score represents an increase of one grade or band in the respective measure, e.g. at A level the difference between being awarded grades BCD and grades BCC. (The increase in

[6] A measure of internal reliability or consistency of the items in a scale. Like other reliability coefficients Cronbach's alpha ranges from 0 to 1. Scores towards the high end of the range suggest that the items in a scale are measuring the same thing.

SAT® score associated with a change in GCSE average points is much larger as it reflects an increase of one grade in a student's average grade.) The significant β coefficients[7] from the regression models in Tables 4.2, 4.3, 5.1 and 5.2 are provided in the appendix to the 2008 report (Kirkup *et al*, 2008).

[7] The β coefficient is a dimensionless quantity showing the strength of the relationship between each variable and the outcome, controlling for all other variables in the analysis.

Table 4.2: Significant predictors of SAT® scores, including IDACI measure

Background variable	Predicted change in SAT® score			
	Overall SAT® score	Reading	Maths	Writing