Raffe, D., Croxford, L., Iannelli, C., Shapira, M. & Howieson, C. (2006). *Social-Class Inequalities in Education in England and Scotland*. Special CES Briefing No. 40. Edinburgh: CES.

Rasbash, J., Browne, W., Healy, M., Cameron, B. & Charlton, C. (2005). *MLwiN version 2.02*. London: Institute of Education.

Rutter, M.L., Maugham, B., Mortimore, P., Ousten, J. & Smith, A. (1979). *Fifteen thousand hours*. London: Open Books.

Snijders, T. and Bosker, R (1999). *Multilevel analysis. An introduction to basic and advanced multilevel modeling*. London: Sage Publications.

Vidal Rodeiro, C.L. (2006). *Factors determining examination success at A-level: a study focussed on A-level Chemistry and A-level Business Studies*. Internal Report. Cambridge: Cambridge Assessment.

EXAMINATIONS RESEARCH

A-level uptake: 'Crunchier subjects' and the 'Cracker effect'

John F. Bell, Eva Malacova, Carmen L. Vidal Rodeiro Research Division
Mark Shannon New Developments

One of the claims made about A-levels is that students are opting for the allegedly easier subjects at A-level. For example, Boris Johnson stated in the *Observer* (July 9, 2006) that 'This year, as every year for the last two decades, we are seeing a drift away from crunchier subjects such as sciences, maths and languages.' More recently, Cambridge University produced a list of A-level subjects that provide a less effective preparation for their courses, for example, Business Studies, Media Studies, and Physical Education, Sports Studies. On their website (http://www.cam.ac.uk/admissions/undergraduate/requirements/), it is stated 'To be a realistic applicant, Cambridge applicants would be expected to have no more than one of these subjects'[1]. It must be stressed that the term 'less effective preparation' refers to the courses offered by what is a highly selective university – these A-levels can be highly relevant and effective preparations for courses offered by other higher education institutions. It is also worth noting that some subjects not on the list had to struggle to gain acceptance. For example, Tillyard (1958) wrote:

> ... [in 1878] it was unthinkable that English should be recognised as an independent study; it could enter Cambridge only on the warrant of a faint respectability reflected from modern languages.

Opponents of English could be quite outspoken, for instance, Edward Augustus Freeman, the Regius Professor of Modern History at Oxford, in a broadside published in 1887 in the London *Times* wrote:

> There are many things fit for a man's personal study, which are not fit for University examinations. One of these is "literature."... [We are told] that it "cultivates the taste, educates the sympathies, enlarges the mind." Excellent results against which no one has a word to say. Only we cannot examine in tastes and sympathies.

As late as 1965, Robson used the first lecture arranged by the F.R. Leavis Lectureship Trust to argue that English Studies met the conventional criteria for admission to a *studium generale*[2]. Also, in 1887 the congregation of Oxford University voted against an Honour School of Modern European Languages. The Warden of All Souls objected because of 'the depreciation and exclusion of Greek and Latin' and that 'it confused the whole conception of academical studies, and dragged

the subjects fit for more advanced years into undergraduate life' (*Times*, 7 November, 1887). However, as Emperor Loathair I (795–855) said 'Tempora mutantur, nos et mutamur in illis'.[3] Whilst it might be possible to idly speculate what a Regius Professor of Media Studies at a 22nd century Cambridge University might make of the current situation, it is probably more informative to consider what exactly is happening with A-levels and determine if the changes are as dramatic as is implied in the media.

In this article we investigate the uptake of A-levels in England. We consider the A-level results for all year 13 students (eighteen-year-olds) in 2001 to 2005 (more detailed analyses for the earlier years can be found in Bell, Malacova and Shannon, 2003, 2005[4]). This period covers the transition to Curriculum 2000 because the new A-levels that were started then were completed in 2002. This reform split A-levels into two. First, a free standing qualification called the Advanced Subsidiary covering the first year of the course was introduced. Secondly, the A-level was obtained by combining results of AS modules with A2 modules. The aim of this reform was that students would study for four or five subjects at AS in the first year of the sixth form and then choose three of them to continue on to A-level. The objective of this reform was to broaden the curriculum and to provide more balance. This is seen as a desirable outcome in many areas of higher education. For example, all medical schools (except Dundee and Edinburgh) encourage potential applicants to take a combination of science and non-science subjects (Clarke, 2005). These medical school policies have implications for the A-level science uptake for the higher attaining candidates. Given that Chemistry is nearly always compulsory and Biology often is, then the effect would be most pronounced in Physics and Mathematics. In general, there are two processes that need to be considered. First, broadening the

---

1. There are exceptions and it is always advisable to check the Cambridge University website for the precise requirement for a course.

2. A recognised university. Originally an institution recognised by the Holy Roman Empire and whose status was confirmed by Papal Bull. Cambridge was formally acknowledged as one in 1290.

3. For those who have not had a classical education: 'Times are a-changing and we change with them.'

4. The analyses differ from those in this article because they include General Studies A-level.

curriculum would lead to a decline in the number of science A-levels as science specialists are encouraged to take other subjects. Secondly, the reverse process is true for non-scientists. In both cases, the change is likely to affect the student's least favourite or least relevant subject. This argument means these changes only affect the pool of qualified candidates actually applying for higher education courses when students change their future plans during their A-level studies, for example, students who would have taken only science A-levels but who substituted Physics with a non-science would have been unlikely to continue with Physics after A-level had they remained science specialists.

Uptake of the most common A-level subjects

In Tables 1–3, the uptake of A-level subjects (strictly, the subject has been defined by the subject code used in the database rather than the specification name) with the highest entries is presented. For clarity, these subjects have been divided into three groups: science and mathematics; arts, languages and sports science; social science and humanities. Candidates were classified by sex and prior attainment at GCSE level. The GCSE grades for the candidates were converted into scores (A*–8, A–7, B–6, etc.) and a mean GCSE score was computed and used to divide the candidates into three attainment groups: low, medium and high. The cut scores were chosen such that they divided the whole A-level entry in three approximately equally sized groups and were carried over for future years. Uptake by attainment is an important issue. Since elite institutions are more likely to require good grades and candidates with higher prior attainment are more likely to obtain such grades, it follows that changes in uptake of subjects by high attaining students can have important implications on the pool of available applicants for courses at these institutions.

Table 1 presents the percentages of A-level students (i.e. having one A-level result) taking each of English and modern language subjects. Changes over the whole period greater than 2 percentage points have been identified in **bold** for declines and *italics* for increases. For all the subjects in this group, uptake is much greater for female students. Uptake also increases with increasing prior attainment for English Literature and the modern languages. There has been a decline in the uptake by female students for English Literature, French and German. This decline is also associated with medium and high prior attainment.

In Table 2 uptakes for arts, media studies and sport studies are presented. The highlighted trends are a decline of females taking Art and Design associated with the change to Curriculum 2000, an increase in Drama uptake, an increase in Media Studies at the time of the change to Curriculum 2000 and an increase in uptake of Sports Studies by candidates with medium prior attainment.

Table 3 is for the humanities and social sciences. Three subjects, Business Studies, Economics, and Geography are all declining except for candidates with high prior attainment. The 'crunchy' subject, History, had an increased uptake by males as did Politics. The increase in Politics was associated with male and high prior attainment candidates. Religious studies was also increasing in popularity. However, the largest changes are associated with Psychology which has the largest increase of any A-level subject, for example, almost one in four females taking three or more A-levels take Psychology. This has been referred to as the 'Cracker[5] effect' because it is argued that uptake has been influenced by the

5. A slang term for a criminal psychologist and used as a name for a popular TV series about one.

Table 1 : Changes in uptake in English and Modern Languages
(% of students with at least one A-level result)

	Year	English	English Language	English Literature	French	German	Spanish
All	2001	7	6	21	7	**4**	2
	2002	6	7	21	6	**3**	2
	2003	6	6	20	6	**3**	2
	2004	6	6	20	6	**2**	2
	2005	6	8	20	6	**2**	2
Male	2001	4	4	13	4	2	1
	2002	4	5	14	4	2	1
	2003	4	5	12	4	2	1
	2004	4	4	12	4	2	2
	2005	4	6	14	4	2	2
Female	2001	9	8	**29**	9	**4**	3
	2002	8	8	**27**	8	**4**	3
	2003	7	8	**26**	7	**3**	3
	2004	8	8	**26**	6	**4**	2
	2005	8	8	**26**	6	**2**	2
Low	2001	8	7	16	2	1	1
	2002	7	7	17	1	1	1
	2003	6	7	15	1	1	1
	2004	6	6	14	2	2	0
	2005	6	8	16	2	2	0
Medium	2001	7	7	**23**	6	3	2
	2002	7	8	**21**	4	2	2
	2003	7	8	**20**	4	2	1
	2004	6	8	**20**	4	2	2
	2005	8	8	**20**	4	2	2
High	2001	5	4	**26**	14	6	4
	2002	6	5	**25**	12	5	4
	2003	5	5	**24**	11	5	4
	2004	4	6	**22**	10	4	4
	2005	4	6	**24**	10	4	4

Table 2 : Changes in uptake of Arts, Media and Sport/PE studies
(% of students with at least one A-level result)

	Year	Art & Des.	Drama	Media/Film/ TV. Stds.	Music	Sport/ P.E.Stds.
All	2001	9	*4*	6	2	7
	2002	10	*6*	8	3	7
	2003	10	*6*	8	3	8
	2004	8	*6*	8	2	8
	2005	8	*6*	8	2	8
Male	2001	7	*2*	5	2	9
	2002	7	*3*	7	2	10
	2003	8	*4*	7	2	11
	2004	8	*4*	8	2	10
	2005	8	*4*	8	2	10
Female	2001	**11**	5	6	3	5
	2002	**8**	8	8	3	5
	2003	**9**	8	9	3	5
	2004	**8**	8	10	2	6
	2005	**8**	8	10	2	6
Low	2001	10	*4*	9	1	10
	2002	11	*7*	12	2	10
	2003	11	*7*	13	1	11
	2004	10	*6*	14	2	10
	2005	10	*6*	14	2	10
Medium	2001	10	*4*	6	2	*8*
	2002	11	*7*	9	2	*9*
	2003	12	*7*	9	2	*10*
	2004	10	*6*	10	2	*10*
	2005	10	*8*	10	2	*10*
High	2001	8	*3*	2	4	*3*
	2002	9	*5*	3	3	*4*
	2003	10	*5*	4	4	*5*
	2004	10	*6*	4	4	*4*
	2005	10	*6*	4	4	*6*

Table 3 : Changes in uptake in Humanities and Social Sciences

(% of students with at least one A-level result)

Group	Year	Busi. St.	Econ.	Geog	History	Law	Politics	Psych.	Relig. Stds.	Socio.
All	2001	14	7	15	15	3	3	10	3	9
	2002	13	6	15	17	4	3	13	4	9
	2003	13	6	14	16	4	4	14	5	9
	2004	12	6	14	16	4	4	16	6	10
	2005	12	6	12	18	6	4	18	6	10
Male	2001	16	10	18	15	3	4	5	2	4
	2002	17	9	17	17	3	4	6	2	4
	2003	16	9	17	17	4	5	7	3	4
	2004	16	8	16	18	4	4	8	4	4
	2005	14	8	14	18	4	6	10	4	4
Female	2001	12	4	13	15	4	3	15	5	13
	2002	10	3	12	16	4	3	18	5	13
	2003	10	3	12	16	5	3	20	6	13
	2004	8	2	12	16	6	2	22	6	12
	2005	8	2	10	16	6	2	24	8	14
Low	2001	17	4	11	10	5	2	11	3	13
	2002	15	3	11	11	5	2	12	3	12
	2003	15	2	10	10	5	2	13	4	12
	2004	14	2	8	10	6	2	14	4	12
	2005	12	2	8	10	6	2	14	4	12
Medium	2001	17	7	18	16	3	3	12	4	10
	2002	16	5	16	16	4	3	16	4	11
	2003	16	5	16	16	5	3	18	5	11
	2004	14	4	14	16	6	4	20	6	12
	2005	14	4	14	16	6	4	22	6	12
High	2001	8	9	17	21	2	4	7	4	4
	2002	9	9	17	22	2	4	11	5	5
	2003	8	9	16	22	3	5	13	5	5
	2004	8	8	16	22	4	4	14	6	6
	2005	8	8	16	22	4	6	14	6	6

Table 4 : Uptake of Science and Mathematics subjects

(% of students with at least one A-level result)

	Year	Biology	Chem.	Com. Stds	D & T design	ICT	Maths	Further Maths	Physics
All	2001	19	16	5	2	3	24	2	13
	2002	19	14	4	6	7	19	2	13
	2003	18	13	4	6	7	19	2	12
	2004	18	14	2	6	6	20	2	10
	2005	18	14	2	6	6	18	2	10
Male	2001	16	18	9	4	4	32	4	22
	2002	16	15	8	9	10	26	3	21
	2003	15	14	7	9	10	26	3	20
	2004	14	14	6	9	8	26	4	18
	2005	16	16	4	9	8	26	4	18
Female	2001	22	14	1	1	2	17	1	5
	2002	22	14	1	4	4	13	1	5
	2003	20	13	1	4	5	13	1	5
	2004	20	12	*	4	4	14	*	4
	2005	20	12	*	4	4	12	*	4
Low	2001	9	5	6	3	4	8	*	5
	2002	7	3	4	8	9	4	*	4
	2003	6	3	3	8	10	4	*	3
	2004	6	4	2	8	8	6	*	4
	2005	6	4	2	8	8	6	*	4
Medium	2001	19	12	6	3	3	21	1	11
	2002	17	10	5	7	8	13	1	10
	2003	16	8	4	8	9	13	1	9
	2004	14	8	4	8	8	12	*	8
	2005	14	8	2	8	6	12	*	8
High	2001	30	30	3	1	1	42	6	22
	2002	30	27	3	4	4	35	4	21
	2003	29	25	3	4	4	34	4	19
	2004	28	26	2	4	4	34	4	18
	2005	30	26	2	4	2	32	4	16

* denotes less than 0.5% uptake

increasing prominence of psychologists in television drama. (http://news.bbc.co.uk/1/hi/education/1635122.stm)

Finally, we consider the uptake of science and mathematics subjects. The longer term trends in Mathematics were considered in Issue 2 of *Research Matters* (Bell, 2006). There have been declines in uptake for females taking Biology and Chemistry. These are particularly associated with medium levels of prior attainment. For Physics the decline also occurs for high attaining candidates. It is important to note that throughout the period under consideration entries in the three traditional sciences and Mathematics have been dominated by candidates with high levels of GCSE attainment. This raises the question as to whether it is desirable for advanced studies in these subjects to be increasingly the preserve of an academic elite. There are clearly issues with the perceived difficulty of these subjects.

It is not enough to consider uptake of individual subjects. Combinations of subjects are also important as indicated by the University of Cambridge's concerns mentioned in the first paragraph of this article. Analysing combinations of subjects is not as straightforward as it seems because there were 23,963 combinations of individual subjects in 2001 (Bell, Malacova and Shannon, 2003, 2005). Therefore, it is necessary to group subjects to analyse combinations.

The first group to be considered is based on the Cambridge list of subjects that are less effective preparation for Cambridge courses (LEPs) (http://www.cam.ac.uk/admissions/undergraduate/requirements/). The subjects that they list are: Accounting, Art and Design, Business Studies, Communication Studies, Dance, Design and Technology, Drama/Theatre Studies, Film Studies, Health and Social Care, Home Economics, Information and Communication Technology, Leisure Studies, Media Studies, Music Technology, Performance Studies, Performing Arts, Photography, Physical Education, Sports Studies, and Travel and Tourism.

It is not the study of individual LEPs that is the perceived problem but rather the studying of too many of them. We decided to investigate the number of these LEPs taken by candidates with three or more A-level results. In Figure 1, 100% stacked area charts are presented for the number of less effective preparation A-levels (LEPs). The data is presented by gender and by prior attainment. The prior attainment is based on mean GCSE with the cut scores that divided the 2001 A-level candidates into three approximately equal groups. The darkest area at the bottom represents the candidates not taking any LEPs. The lighter grey area represents those taking one LEP and the next two areas 2 and 3+ LEPs (3+ is the top area). The data for all the tables in this report can be found in Vidal Rodeiro (2006) on the Cambridge Assessment website. http://www.cambridgeassessment.org.uk/research/statisticalreports/

It is clear that the majority of students still take at least two 'acceptable' subjects. Male students are less likely to take LEPs than female students. There is an interesting effect relating to Curriculum 2000. In the first year, there was a decrease in LEP subjects by low prior attainment candidates and the reverse pattern for high prior attainment candidates. It is likely that pattern was based on centres gaining experience of the new specification (in particular the A2 modules) and the attitudes of universities to certain subjects.

It should be recognised that the majority of candidates for Cambridge and other elite institutions likely to have similar restrictions on suitability of A-levels will largely recruit from candidates with high prior GCSE attainment. Only 5% of the candidates in 2001 made a choice of A-levels that included 2 or more LEPs. This increased to 17% in 2002 and then fell to about 9% for the remaining three years (note that for candidates performing at the level of the vast majority of successful Cambridge applicants, i.e. those with three grade As, the percentages affected are much smaller.)

Figure 1 : Uptake of 'Less Effective Preparation' A-levels   KEY: ■ 0 ▨ 1 ▥ 2 ☐ 3+

(a) Female   (b) Male   (c) All

(d) Low   (e) Medium   (f) High
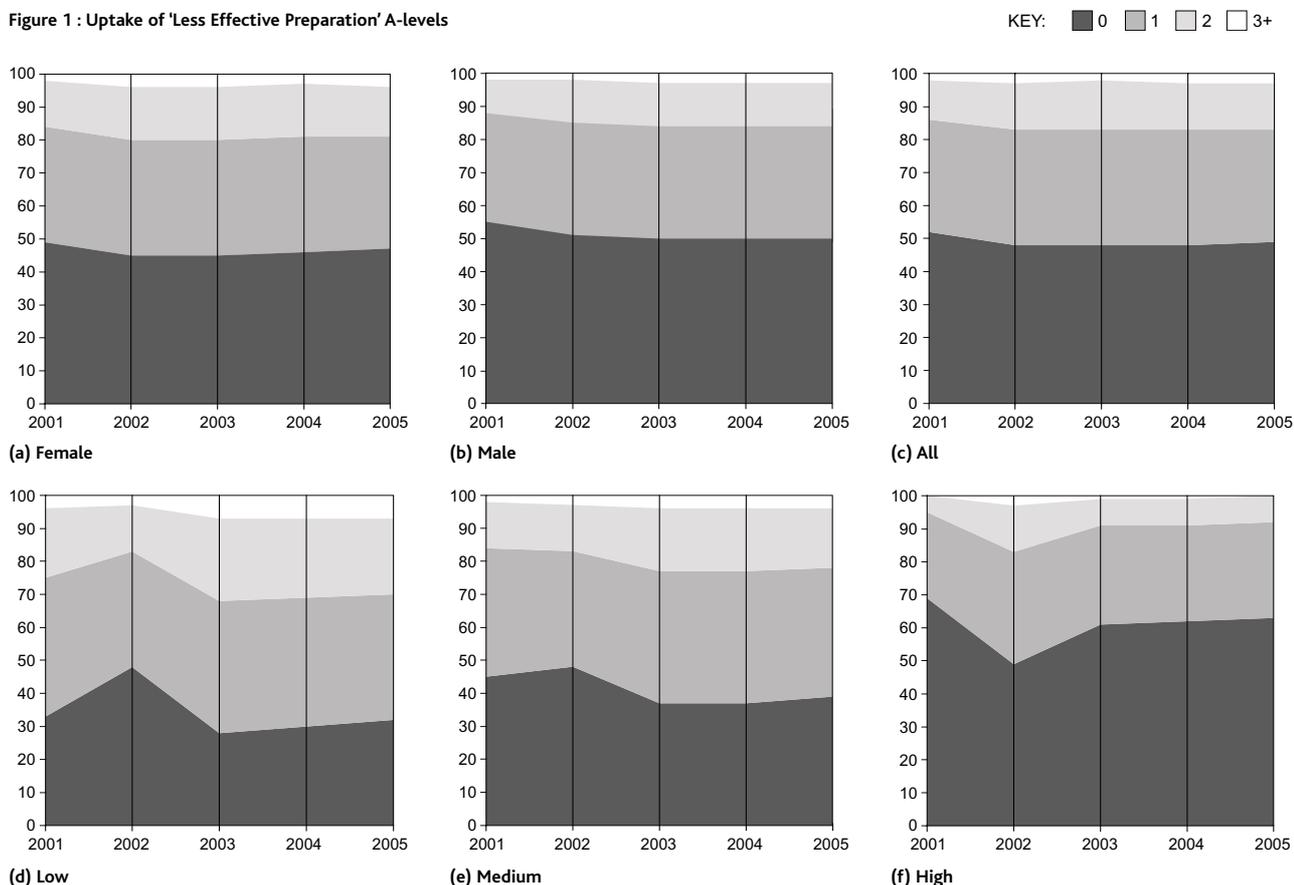
One of the aims of Curriculum 2000 was to broaden students' experiences and to discourage early specialisation. In Figure 2, the uptake of five subject areas at A-level is presented. The five areas were: Science/Mathematics, English, Languages, Social Science/Humanities and Arts. Grouping subjects is not a straightforward task and the allocation of subject areas is always debatable (at the time of analysis psychology specifications were usually grouped with the social sciences but modified specifications starting in 2008 are going to be classified as sciences). Some subjects do not necessarily fit comfortably in any category. More details of the subject areas can be found in Bell, Shannon and Malacova. (2003, 2005). The categories were originally derived to illustrate how close the current situation is to a balanced diploma based on existing A-levels. The percentages in the figure relate to the number of A-level students taking at least one of the subjects in the subject area in the population of students taking at least three A-levels.

The most obvious feature of Figure 2 is the stability of the uptake for most subject areas. The clearest trend is the decline in the number of students taking Modern Languages where the rate of decline is faster for female students. There were no consistent trends for the other domains.

There are however large differences between the subject areas. For female students, the subject area with the highest uptake is Social Science/Humanities (note this is made up of many more A-level subjects than the other areas). The remaining subject areas in descending order of uptake are English, Science/Mathematics, Arts and Modern Languages. The pattern for male students is different. The uptake of the Science/Mathematics subject area is similar to that of the Social Science/Humanities area. The uptakes of subjects in the English, Arts and the Modern Languages groups are much lower than for females.

Figure 2 also shows that the uptake of subject domains is related to ability. Uptake of Arts, English, and Social Science domains all decline

with increasing prior attainment. This is most marked for the Arts domain where the percentage uptake is approximately halved. Uptake for the remaining two domains, Modern Languages and Sciences, increases with prior attainment. This relationship is strongest for Modern Languages with approximately one in twenty students in the lowest prior attainment group taking at least one modern language compared with one in five for the high prior attainment group.

John Dunford, general secretary of the Association of School and College Leaders stated that Modern Languages were in freefall (*Guardian*, 24 August, 2006). Whilst the decline has been substantial, there are two features of it that are interesting. First, it is much smaller for males compared with females and, secondly, for low attaining candidates the decline is also smaller. There is a need for further research in these areas.

For the final analyses in this article, subjects were grouped into three different domains; Science and Mathematics, Arts and Languages, Social Science and Humanities (the subject domains involved merging the subject areas of Art, English and Languages into one domain. This categorisation was used in Bell *et al*. (2005) to investigate whether A-level subject choice was balanced). Using these three domains, it possible to classify candidates taking three or more A-levels into seven groups:

YNN Science/Mathematics only.

NYN Arts /Languages only

NNY Social Sciences/Humanities only

YYN Science/Mathematics and Arts/Languages

YNY Science/Mathematics and Social Sciences/Humanities

NYY Arts / Languages and Social Sciences / Humanities

YYY All three domains

Figure 2 : Uptake of A-level subject groups   KEY: ◆ Science/Maths   □ English   △ Languages   ✕ Social Science/Humanities   ✳ Arts
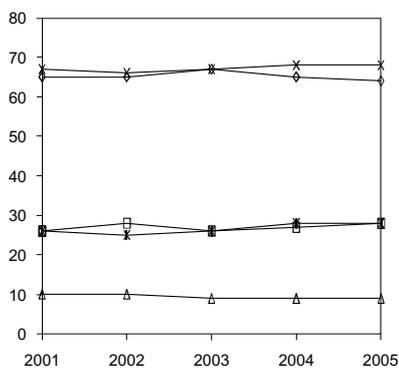
(a) Male
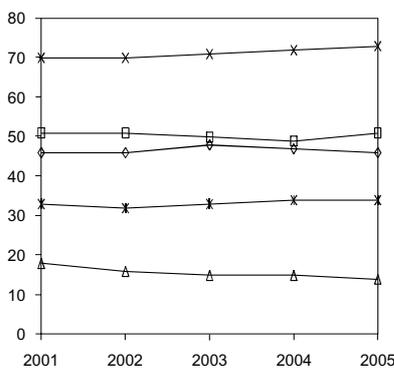
(b) Female
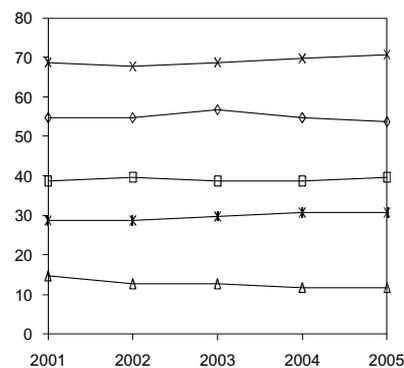
(c) All
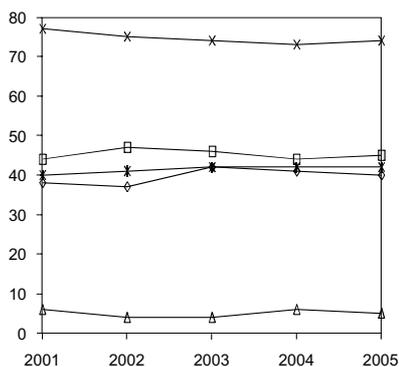
(d) Low
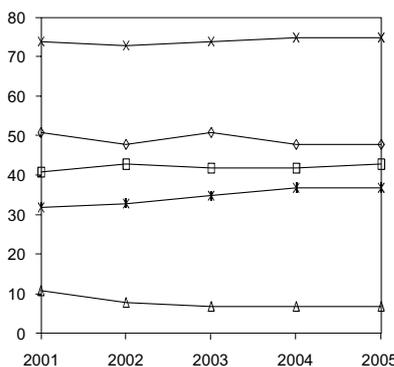
(e) Medium
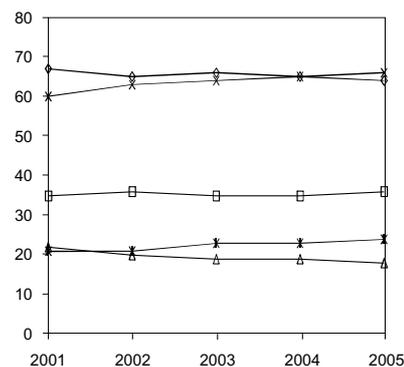
(f) High

Figure 3 presents the percentage uptake of these categories by sex and attainment for 2001 to 2005 for those candidates taking at least three A-levels. In the figures solid lines have been used to denote combinations including science and mathematics and dotted lines for those that do not. Looking at all the panels of Figure 3 it can be seen that for most combinations of domains the changes have been small. The only large changes tend to be associated with science specialists. The increases tend to be smaller and shared between combinations involving science. For the combinations of domains not including Science/Mathematics, there are only small, inconsistent year on year changes. This means that the net effect of broadening the curriculum has been to reduce the amount of science that science specialists study but this has not been matched by an increased uptake by non-scientists.

When all the data presented here are considered in their entirety, there are some noticeable results. First, for most subjects and groups of subjects there has been very little change. For some subjects and groups of subjects there have been changes associated with Curriculum 2000 but the uptakes have subsequently stabilised. Of greater concern are the subjects that have declined through the whole period, for example, Geography, Physics and Modern Languages as a group.

Although there has been a decline in numbers taking general qualifications (GCSE and GCE) that assess modern languages, there is an important development that seems to provide a promising solution. This summer's GCSE results showed a big decline in the number of pupils studying modern foreign languages. French and German suffered the biggest falls in candidates of any subject, with declines of 13.2% and 14.2%. There are, however, alternatives to existing qualifications that may be of use in increasing the number of linguists. Cambridge Assessment has developed a new qualification scheme called Asset

Languages. This is part of the DFES National Languages Strategy (http://www.assetlanguages.org.uk/). They use a 'ladder' of courses similar to music grades and aim to make language learning accessible. More than a quarter of state secondary schools are going to use these qualifications from September 2006. There are also 120 primary schools involved in the scheme. Experience from the first full year of the scheme suggests that it is successful in motivating students' language learning.

For Science and Mathematics, there is a need to consider how these subjects are extended beyond a very able elite. When considering trends in uptake, a common mistake is to use what was described in the TV series *Yes Minister* as the politician's syllogism: 'Something must be done. This is something. Therefore, this must be done'. Before acting it is better to gain an understanding of the underlying causes for the trend. This article is only the first step in understanding uptake in A-levels. Before acting it is necessary to understand the processes that have led to the situation described in this article. This requires the collection of additional information. For this reason, Cambridge Assessment is currently conducting a large scale survey (with the Association of Colleges) investigating why students choose particular A-levels.

Acknowledgements

References

Bell, J.F. (2006). The curious case of the disappearing mathematicians. *Research Matters: A Cambridge Assessment Publication*, **2**, 21–23.

Bell, J.F., Malacova, E. & Shannon, M. (2003). *The Changing Pattern of A-level/AS uptake in England*. A paper presented at the British Educational Research

Figure 3 : Uptake of combinations of subject domains   KEY: ◆ YNN   ⬚ NYN   △ NNY   ✳ YYN   ✲ YNY   ○ NYY   ◇ YYY

(a) Male

(b) Female

(c) All

(d) Low

(e) Medium

(f) High
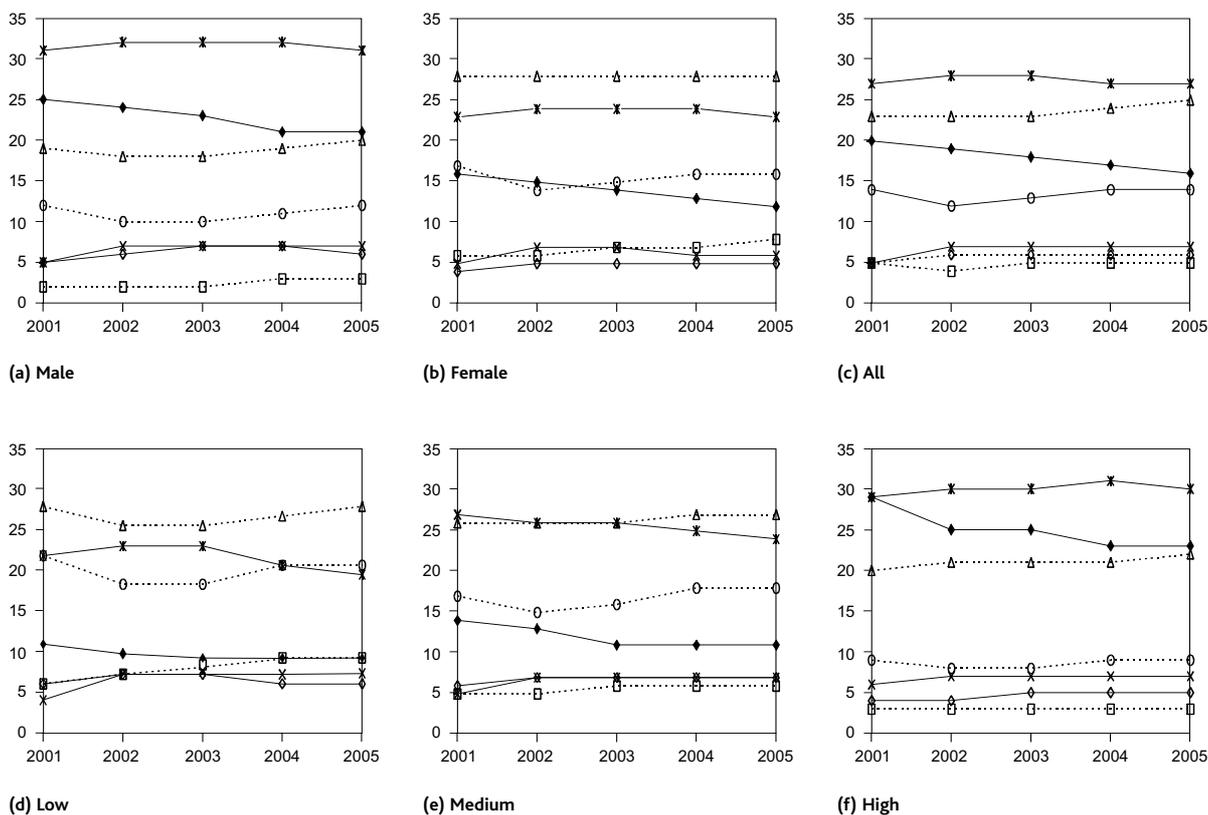
Association Annual Conference, Edinburgh, September 2003.
http://www.cambridgeassessment.org.uk/research/confproceedingsetc/BERA2
003JBEMMS/file/

Bell, J.F., Malacova, E. & Shannon, M. (2005). The changing pattern of A level/AS
uptake in England. *The Curriculum Journal*, **16**, 3, 391–400.

Clarke, P. (2005). *Curriculum 2000 and other qualifications. A survey of UK medical
schools' requirements and selection policies*. Cardiff University.
http://www.ucas.ac.uk/candq/curr2000/medical.pdf

Centre for Education and Industry (2003). *Widening participation in the physical
sciences: An investigation into factor influencing the uptake of physics and
chemistry*. University of Warwick, Coventry: Centre for Education and Industry.

Robson, W. W. (1965). *English as a university subject*. London: Cambridge
University Press.

Smithers, A. & Robinson, P. (2006). *Physics in schools and universities: II. Patterns
and policies*. Buckingham: Centre for Education and Employment Research.
http://www.buckingham.ac.uk/education/research/ceer/pdfs/physicsprint-
2.pdf

Tillyard, E.M.W. (1958). *The muse unchained: An intimate account of the
revolution in English Studies at Cambridge*. London: Bowes and Bowes.

Vidal Rodeiro, C.L. (2006). Uptake of GCE A-level subjects in England 2001–2005.
Statistical Report Series No. 3.
http://www.cambridgeassessment.org.uk/research/statisticalreports/

Discussion piece: The psychometric principles of assessment

Professor John Rust Psychometrics Centre

Psychometrics is the science of psychological assessment, and is a foundation of assessment and measurement. Within psychometrics there are four fundamental principles whereby the quality of an assessment is judged. These are (1) reliability, (2) validity, (3) standardisation and (4) freedom from bias. Reliability is the extent to which an assessment is free from error; validity is the extent to which a test or examination assesses what it purports to assess; standardisation gives us information on how the result of an assessment is to be judged, and freedom from bias examines the extent and causes of differences between groups. These four principles inform not only test use but also the entire process of test development, from the original curriculum or job specification, via the choice and appraisal of examination questions and test items, through to the eventual evaluation of the success or otherwise of the assessment itself.

No assessment can be perfectly reliable, and this applies not only to the measurements we make in education or psychology, but to all types of measurement. Measurements range in accuracy from the exceptionally high levels now obtained for the speed of light and the time of day, through measurements of length and area used in surveying, to the lower levels attainable for measurement of blood pressure and haematological assays used in medicine, to the tests of ability, achievement and character with which we are familiar in the education and recruitment testing world. Hence, in all these cases our expectations are different. Reliability is assessed on a scale of zero to one, with a score of 0.00 indicating no reliability at all, and a score of 1.00 representing perfect reliability. Over a century of human testing has shown us that we can expect reliabilities ranging from 0.95 for a very carefully constructed and individually assessed test of ability, through 0.85 for group tests of ability; about 0.75 for personality tests; 0.5 for rating scales and down to about 0.2 or 0.3 for projective tests or tests of creativity.

There are several ways in which reliability can be assessed and most of them involve making multiple measurements. Inter-rater reliability is the extent to which examiners agree or disagree about the mark that a candidate should be given when the assessments are made independently. This is the most relevant form of reliability for many forms of school examinations, and the assessment of coursework or essays has an apparent upper limit of about 0.65. It is interesting that in spite of frequent attempts to improve on inter-relater reliability of examiners, for example by setting detailed marking criteria, it is unusual to find a value that goes much above this figure. Constraining the judgement of individual examiners can, if taken to extremes, lead to just another multiple choice test by another name.

Whatever efforts are put into improving the agreement between markers, it is not the only source of unreliability and may not even be the most important. In multiple choice examinations, for example, the inter-rater reliability is often as high as 0.99, because the only disagreement between raters is in reading the scores from the answer sheets. However, unreliability still arises for many other reasons, such as the state of the candidate (tired, ill, anxious etc.) the environment in which the test is taken, events at home or among the peer group, or the concordance between the content of the examination and the revision strategy used by the candidate, to name but a few. These forms of reliability are called test stability, and one way of obtaining this is by administering the same test or examination to the same group of individuals on two or more occasions and comparing the results. But this can only be an estimate, as the prior experience of having sat the same examination will tend to affect the second sitting in some way. In spite of this it is essential that we have some way of estimating stability effects for our assessments. Given all the possible sources of instability, we expect an upper limit of at most about 0.85 on the expected reliability of a multiple choice school examination. It is important to note I am not here trying to do full justice to issues of reliability, I am illustrating the importance of the application of psychometric principles.

Reliability in assessment is just the first step, however. A score can be perfectly reliable and still utterly invalid for a particular application. Astrological charts or diagnoses made on the basis of graphology (handwriting analysis) may be very reliable in that astrologers or graphologists using the same system will usually be in agreement about