Learning in practice

Intellectual aptitude tests and A levels for selecting UK school leaver entrants for medical school

I C McManus, David A Powis, Richard Wakeford, Eamonn Ferguson, David James, Peter Richards

An extension of A level grades is the most promising alternative to intellectual aptitude tests for selecting students for medical school

How to make the selection of medical students effective, fair, and open has been contentious for many years.[w1] A levels are a major component of selection for entry of school-leavers into UK universities and medical schools,[w2] but intellectual aptitude tests for the selection of medical students are burgeoning—they include the Oxford medicine admissions test[1] and the Australian graduate medical school admissions test [2] (table). Tests such as the thinking skills assessment[w3] are being promoted for student selection generally. The reasons include a political climate in which government ministers are advocating alternatives to A levels, some support for them in the Schwartz report on admissions to higher education,[3] lobbying from organisations such as the Sutton Trust,[w4] and the difficulty of distinguishing between the growing numbers of students achieving three A grades at A level. We examine the problems that intellectual aptitude tests are addressing, their drawbacks, any evidence that they are helpful, and alternatives.

Medical schools need selection procedures that are evidence based and legally defensible. We therefore explored a series of questions around these developments.

Many UK school leavers apply to university with other educational qualifications, including the international Baccalaureat and Scottish highers. Medical schools are increasingly admitting entrants other than directly from school. These different mechanisms of entry require separate study. We discuss the main school-leaver route through A levels.

Summary points

So many applicants are achieving top grades at A level that it is increasingly impractical to select for medical schools primarily on such achievement

Schools are introducing tests of intellectual aptitude without evidence of appropriateness, accuracy, or added value, making them open to legal challenge

Since the 1970s, university achievement has been shown to be predicted by A levels but not by intelligence tests

The discriminative ability of A levels might be restored by introducing A+ and A++ grades, as recommended in the Tomlinson Report

If additional grades at A level cannot be introduced, medical schools could collectively commission and validate a new test of high grade scientific knowledge and understanding

An argument exists for also developing and validating tests of non-cognitive variables in selection, including interpersonal communication skills, motivation, and probity

Are A levels predictive of outcome?

Many beliefs are strongly held about undergraduate student selection but without any "visible means of support"[4]: one is that A levels are not predictive of outcome at university. The opposite is true. A study of 79 005 18 year old students entering university in 1997-8 and followed through until 2000-1 (www.hefce.ac.uk) shows a clear relation between A level grades and university outcome (fig 1). The result is compatible with many other studies of students in general,[5w5 w6] and of medical students in particular.[5–8w7-w10] Small studies of individual students in individual years at individual institutions are unlikely to find such correlations—the reasons being statistical, including lack of power, restriction of range, and attenuation of correlations caused by unreliability of outcome measures (see bmj.com).

Why are A levels predictive?

The three broad reasons why A levels may predict outcome in medicine are: cognitive ability—A levels are indirect measures of intelligence, and intelligence

P+ *Supplementary information and figures are on bmj.com*

Department of Psychology, University College London, London WC1E 6BT
I C McManus *professor of psychology and medical education*
University of Newcastle, New South Wales 2308, Australia
David A Powis *professor of health professional education*
Cambridge University School of Clinical Medicine, Addenbrooke's Hospital, Cambridge CB2 2QQ
Richard Wakeford *educational adviser*
School of Psychology, University of Nottingham, Nottingham NG7 2RD
Eamonn Ferguson *professor of health psychology*
Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH
David James *professor of fetomaternal medicine*
Hughes Hall, University of Cambridge, Cambridge CB1 2EW
Peter Richards *president*
Correspondence to: I C McManus i.mcmanus@ucl.ac.uk

BMJ 2005;331:555–60

Aptitude tests currently used in United Kingdom by medical schools and other university courses

Test	Name	Further information	Comments
BMAT	Biomedical admissions test	www.bmat.org.uk	Used by Cambridge, Imperial College, Oxford, and University College London, as well as three veterinary schools
GAMSAT	Graduate medical school admissions test	www.acer.edu.au/gamsat	Used for selection by Australian graduate medical schools. At present it is used by four graduate entry schools in United Kingdom. UK version: GAMSAT UK
MSAT	Medical school admissions test	www.acer.edu.au/msat	Used by three UK medical schools
MVAT	Medical and veterinary admissions test	No published details	Developed in Cambridge and was a precursor to the biomedical admissions test
OMAT	Oxford medicine admissions test	See reference 3	Developed in Oxford and was a precursor to the biomedical admissions test
PQA	Personal qualities assessment	www.pqa.net.au	Subtests of mental agility, interpersonal values, and interpersonal traits. Administered in several UK medical schools on a research basis only
TSA	Thinking skills assessment	www.cam.ac.uk/cambuniv/undergrad/tests/tsa.html	Used by several Cambridge colleges for selection in a range of disciplines, of which computer science is presently the predominant one

correlates with many biological and social outcomes[9]; substantive content—A levels provide students with a broad array of facts, ideas, and theories about disciplines such as biology and chemistry, which provide a necessary conceptual underpinning for medicine; and motivation and personality—achieving high grades at A level requires appropriate motivation, commitment, personality, and attitudes, traits that are also beneficial at medical school and for lifelong learning.

Cognitive ability alone cannot be the main basis of the predictive ability of A levels, because measures of intelligence and intellectual aptitude alone are poor predictors of performance at university. This is not surprising, as an old axiom of psychology says that "the best predictor of future behaviour is past behaviour," here meaning that the future progress in passing medical school examinations will best be predicted by performance in past examinations. Of course success in medicine and being a good doctor are not identical nor are either of these the same as passing medical school examinations; but those who fail medical school examinations and have to leave medical school never become doctors of any sort. The predictive value of A levels most likely results either from their substantive content, their surrogate assessment of motivation to succeed, or both.

Separating the substantive and motivational components of A levels is straightforward in principle. If the substantive content of A levels is important for prediction then there will be a better prediction of outcome from disciplines underpinning medical science, such as biology and chemistry, than there will be from other subjects—for example, music, French, or economics. Alternatively if motivational factors are the main basis for the predictive power of A levels, indicating pertinent personality traits and attitudes such as commitment, the particular subject taken will be less relevant and an A grade in music, French, or economics will be found to predict performance at medical school as well as an A grade in biology or chemistry. Few analyses, however, differentiate between these factors. But evidence is increasing that A level chemistry is a particularly good predictor of performance in basic medical science examinations[6 10] (although not all studies find an effect[w7 w11]), and A level biology also seems to be important.[6 11] As almost all medical school entrants take at least two science A levels, however, one of which is chemistry, this leaves little variance to partition. It should also be remembered that there are probably true differences in the difficulty of A levels, with A grades easier to achieve in subjects such as photography, art, Italian, and business studies, than in chemistry, physics, Latin, French, mathematics, biology, and German.[12] A clear test of the need for substantive content will occur should medical schools choose to admit students without A level chemistry, as has been suggested,[w12] or perhaps without any science subjects. If students with only arts A levels perform as well as those with science A levels, then the substantive content of A levels is unimportant—and there are suggestions that arts and humanities subjects independently predict outcome.[13 w13]

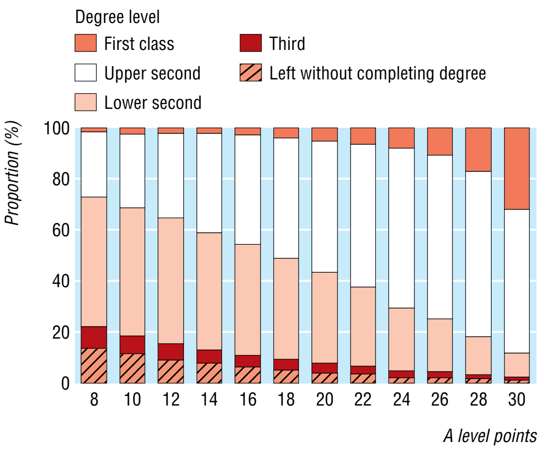


Fig 1 Outcome of students in relation to A level grades in all subjects. Grades are based on best three. A=10, B=8, C=6, D=4, and E=2 points

What do intellectual aptitude tests do?

Aptitude has many meanings,[14] but the glossary to the Schwartz report says that aptitude tests are "designed to measure intellectual capabilities for thinking and reasoning, particularly logical and analytical reasoning abilities."[3] Aptitude, however, also refers to non-cognitive abilities, such as personality. We therefore talk here of "intellectual aptitude," both in the sense described by Schwartz and in the meaning used in the United States for what used to be known as scholastic aptitude tests (SATs; but now the "A" stands for assessment) and which are largely assessments of intellectual ability.[w14]

Most intellectual aptitude tests assess a mixture of what psychologists call fluid intelligence (logic and critical reasoning, or intelligence as process[15]) and crystallised intelligence (or intelligence as knowledge, consisting of general culturally acquired knowledge of vocabulary, geography, and so on). We thus consider intelligence tests and intellectual aptitude tests as broadly equivalent, but distinguish them from achievement tests, such as A levels, which assess knowledge of the content and ideas of academic subjects such as chemistry and mathematics.

Although purveyors of tests such as the biomedical admissions test (table) argue that they are not measures of intelligence but of "critical thinking," there is little agreement on what critical thinking means.[w15-w17] Critical thinking is related more to aspects of normal personality than it is to IQ and reflects a mixture of cognitive abilities (for example, deductive reasoning) and dispositional attitudes to learning (for example, open mindedness).[16] Evidence also shows that critical thinking skills, not dispositions, predict success in examinations[17] and that critical thinking may lack temporal stability.[18] The content and the timed nature of the biomedical admissions test suggest it will correlate highly with conventional IQ tests.

What do intellectual aptitude tests predict?

We know of three studies that have compared intellectual aptitude tests with A levels (see bmj.com).

The investigation into supplementary predictive information for university admissions (ISPIUA) project[5] studied 27 315 sixth formers in 1967, who were given a three hour test of academic aptitude (TAA); 7080 entered university in 1968 and were followed up until 1971. The results were clear: "TAA appears to add little predictive information to that already provided by GCE results [A levels and O levels] and school assessment in general."

The Westminster study[14] followed up 511 entrants to the Westminster Medical School between 1975 and 1982 who had been given a timed IQ test, the AH5.[w18] Intellectual aptitude did not predict outcomes measured in 2002, whereas A level grades were predictive of both academic and career outcomes.

The 1991 cohort study looked at 6901 applicants to UK medical schools in 1990, of whom 3333 were admitted[w19 w20] and followed up.[w21 w22] An abbreviated version of the timed IQ test was given to 786 interviewees.[w18] A levels were predictive of performance in basic medical science examinations, in final clinical examinations, and in part 1 of a postgraduate examination, whereas the intellectual aptitude test was not predictive (see bmj.com).

In the United States (presently outside the hegemony of the A level system) a recent study of dental school admissions[19] evaluated an aptitude selection test, carefully founded in the psychology of cognitive abilities and skill acquisition. The scores were of no predictive value for clinical achievement at the end of the course.

Do aptitude tests add anything to what A levels already tell us?

In interpreting the validity of aptitude tests it should be acknowledged that some aptitude tests are not content free and do assess substantive components. For instance, the biomedical admissions test seeks to test aptitude and skills (problem solving, understanding argument, data analysis, and inference) in section 1 and scientific knowledge and applications in section 2. Section 2 contains questions on biology, chemistry, physics, and mathematics (www.bmat.org.uk) at the level of key stage 4 of the UK national curriculum (www.nc.uk.net). Because most applicants for medical school are studying many of those subjects for A level, section 2 may be a better predictor of university outcome than is section 1, because it is indirectly assessing breadth of scientific background (and in all likelihood it will be found to correlate with A level grades, but not necessarily to add any predictive value).

On balance, A levels predict university achievement mainly because they measure the knowledge and ideas that provide the conceptual scaffolding necessary for building the more advanced study of medicine.[20 w23] As with building a house, the scaffolding will later be taken down and its existence forgotten, but it will nevertheless have played a key part in construction. Motivation and particular personality traits may also be necessary, just as a house is built better and faster by an efficient, conscientious builder. However, intellectual aptitude tests assess neither the fundamental scientific knowledge needed to study medicine nor the motivation and personality traits. Pure intellectual aptitude tests only assess fluid intelligence, and empirically that is a weak predictor of university performance. Intelligence in a builder, although highly desirable, is no substitute for a professional knowledge of the craft of house building.

What are the problems of using A levels for selection?

Reported problems in using A levels in selection are threefold: the increasing numbers of candidates with three A grades at A level; social exclusion, and type of schooling.

A continuing concern of the UK government is that entry to medical school is socially exclusive.[21] The class distribution of entrants has been unchanged for over half a century, with a preponderance of applicants and entrants from social class 1.[w24] It is not clear whether the entrant profile reflects bias by selectors,[w25] an active choice not to apply for medicine by those from social classes IV and V,[w26] or an underlying distribution of ability by social class.[w27]

Although most children in the United Kingdom attend state schools, the minority attending private (independent) schools are over-represented among university entrants, possibly because as a group they achieve higher A level scores.

Can these problems be fixed?

The increasing numbers of candidates with three grade As at A level

Intellectual aptitude tests are seen as a way to stretch the range, continuing to differentiate when A level grades are at their ceiling. The problem is that although these tests undoubtedly provide variance (as indeed would random numbers), it is not useful variance since it does not seem to predict performance at university.

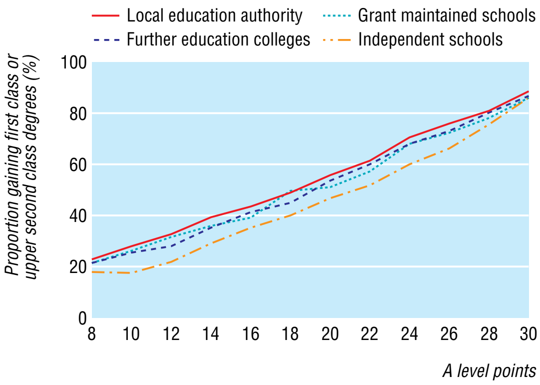


Fig 2 Predictive value for gaining first class or upper second class degree of A levels obtained at different types of school (see bmj.com for definition of school types)

The simplest solution to the ceiling problem is that suggested in the Tomlinson Report[22] of introducing A+ and A++ grades at A level, so that A levels continue to do what they do well, rather than being abandoned and replaced by tests with unproved validity. An appropriate alternative strategy would be to commission a new test of high level scientific knowledge and understanding that measures above the top end of A levels, but it may be better to stay with what we already know.

Social exclusion
It has been suggested that a pool of talented individuals capable of becoming good doctors is excluded by current admission methods. Even if that were so (and we know of no evidence to support it) there is no basis for believing that intellectual aptitude tests are capable of identifying them.

Type of schooling
To tackle possibly unfair over-representation of entrants from independent schools, a case has been made for university selectors taking into account type of school and the relative performance of the school: achieving three grade Cs at A level in a school where most pupils gain three grade Es may predict university achievement better than gaining three grade Cs in a school where most pupils gain three grade Bs. Detailed analyses by the Higher Education Funding Council for England, however, show that after taking into account the grades achieved by an individual student, the aggregate achievement of the school from which the student has come provides no additional prediction of university outcome (see bmj.com). The funding council has good evidence to show that on aggregate, pupils from independent schools under-perform at university compared with those with the same grades from state schools (fig 2).

Intellectual aptitude tests are not a solution to this problem. A solution might be to upgrade the A level grades of applicants from state schools so that, say, one A grade and two B grades are treated as equivalent to two A grades and one B grade from an independent school applicant (that is, increasing by 20 points on the new tariff of the Universities and Colleges Admissions Service (www.ucas), or by 2 points on the older scheme shown in figure 2). Any system should also take into account that many pupils are at independent schools until age 16 and then transfer to (state) sixth form colleges for their A levels. A proper, holistic assessment of each student will, however, require more information than is readily available on the present admissions form.

What is the potential value of non-cognitive aptitude tests?

The aptitude tests we have considered are those that assess cognitive skills. Other skills, however, are needed by doctors, such as the ability to communicate and to empathise, having appropriate motor and sensory skills, and having the appropriate attitudes and ethical standards necessary for professionalism (see for example, www.pqa.net.au). None of these is disputed, and there are strong arguments that selectors can and should take such measures into account at the same time as they are assessing the intellectual skills necessary for coping with the course.[w28 w29] We have not considered such aptitudes in detail here because few UK medical schools are as yet using them in selection (as opposed to research and validation), although initiatives are in progress, both in the United Kingdom and elsewhere.[w30 w31] Situational selection tests have been used in five Flemish medical and dental schools and were found to predict performance in the final first year's examinations better than tests of cognitive ability.[23]

It is also the case, however, that if selection is to be made on the basis of several independent characteristics, then the extent of selection on each is inevitably lower than if there is selection only on any one of them,[24] until eventually the selective power of each is so reduced that "if you select on everything you are actually selecting on nothing."[w32] An attractive argument is that if most students with three grade As at A level (or indeed even lower grades) can cope with a medical course, then instead of looking for selection using A+ and A++ grades, medical schools should be selecting from the existing pool more systematically on non-cognitive measures. That will require validation of the measures, but it might result in a cohort of doctors who are not only academically able but also well suited to medicine because of temperament and attitude. Whether the slight potential decrease in the academic qualifications of entrants will be offset by their increased suitability on non-cognitive measures will depend on a precise knowledge of the relation between academic ability, suitability, and examination performance, and in particular whether these are linear or non-linear and show thresholds. It is an important question that must be answered by empirical study.

Conclusions
Schwartz urged universities to use "tests and approaches that have already been shown to predict undergraduate success" and to assess applicants holistically. We conclude that A levels, using a more finely developed marking system at the top end (A+ and A++ grades, for example) have the greatest potential towards enabling enhanced selection by medical schools' admissions staff: such grades will be maximally robust, in view of the testing time (and coursework) involved.

We understand why the new intellectual aptitude tests are being introduced, but are concerned that they are being introduced uncritically and without published evidence on their reliability and validity. Typically, they involve only an hour or two of testing time and are thus unlikely to have high reliability or generalisability (particularly owing to content specificity), although no data have been published. Their validity can be doubted for good reason, as published studies have found that intellectual aptitude compares poorly with A levels in predicting the outcome of university and medical school, and it has not been shown to add value to the selection process.

The appropriate alternative to refining A level grades would be for the medical schools to commission a new test, reliably assessing high grade scientific knowledge and understanding. At the same time, more research into the value of non-cognitive tests is clearly important and required.

We accept that our criticism of intellectual aptitude tests could be shown to be misplaced when the medical schools using them publish their evidence on predictive validity and reliability. Currently the tests are being justified, not by means of any reported data but by general assertions of organisational quality, unspecified relations between scores and university examinations, and by the observation that admissions staff are using them.[25] Without evidence, medical schools using these tests are vulnerable to legal challenge.

We thank the referees and William James and Chris Whetton for help with preparing this paper.

Contributors: DAP and ICM conceived the article. ICM wrote the first draft. DAP, RW, EF, DJ, and PR provided input. RW revised the draft, following referees' comments. ICM is guarantor.

Funding: None.

Competing interests: DAP is a member of the team developing the personal qualities assessment.

1 James W, Hawkins C. Assessing potential: the development of selection procedures for the Oxford medical course. *Oxford Rev Educ* 2004;30:241-55.
2 Brown P. Admissions testing for entry to medicine. 01: the newsletter of LTSN-01 2004;No 6:27. www.ltsn-01.ac.uk (accessed 6 Jan 2005).
3 Admissions to Higher Education Steering Group. Fair admissions to higher education: recommendations for good practice. Nottingham: Department for Education and Skills Publications, 2004. www.admissions-review.org.uk (accessed 6 Jan 2005).
4 Bekhradnia B, Thompson J. Who does best at university? London: Higher Education Funding Council England, 2002. www.hefce.ac.uk/Learning/whodoes (accessed 6 Jan 2005).
5 Choppin BHL, Orr L, Kurle SDM, Fara P, James G. *The prediction of academic success.* Slough: NFER Publishing, 1973.
6 James D, Chilvers C. Academic and non-academic predictors of success on the Nottingham undergraduate medical course 1970-1995. *Med Educ* 2001;35:1056-64.
7 Lumb AB, Vail A. Comparison of academic, application form and social factors in predicting early performance on the medical course. *Med Educ* 2004;38:1002-5.
8 Ferguson E, James D, Madeley L. Factors associated with success in medical school and in a medical career: systematic review of the literature. *BMJ* 2002;324:952-7.
9 Deary IJ. Differences in mental abilities. *BMJ* 1998;317:1701-3.
10 Montague W, Odds FC. Academic selection criteria and subsequent performance. *Med Educ* 1990;24:151-7.
11 McManus IC, Richards P. Prospective survey of performance of medical students during preclinical years. *BMJ* 1986;293:124-7.
12 Curriculum, evaluation, and management centre. A-level update. www.cemcentre.org/recenttopics/alevelupdate/default.asp (accessed 6 Jan 2005).
13 Grey MR, Pearson SA, Rolfe IE, Kay FJ, Powis DA. How do Australian doctors with different pre-medical school backgrounds perform as interns? *Educ Health* 2001;14:87-96.
14 McManus IC, Smithers E, Partridge P, Keeling A, Fleming PR. A levels and intelligence as predictors of medical careers in UK doctors: 20 year prospective study. *BMJ* 2003;327:139-42.
15 Ackerman PL. A theory of adult intellectual development: process, personality, interests and knowledge. *Intelligence* 1996;22:227-57.
16 Perkins D, Tishman S, Ritchart R, Donis K, Andrade A. Intelligence in the wild: a dispositional view of intellectual traits. *Educ Psychol Rev* 2000;12:269-93.
17 Giddens J, Gloackner GW. The relationship of critical thinking to performance on the NCLEX-RN. *J Nurs Educ* 2005;44:85-9.
18 Pithers RT, Soden R. Critical thinking in education: a review. *Educ Res* 2000;42:237-49.
19 Gray SA, Deem LP, Straja SR. Are traditional cognitive tests useful in predicting clinical success? *J Dent Educ* 2002;66:1241-5.
20 Ferguson E, James D, O'Hehir F, Sanders A. A pilot study of the roles of personality, references and personal statements in relation to performance over the 5 years of a medical degree. *BMJ* 326;2003:429-31.
21 Department of Health. *Medical schools: delivering the doctors of the future.* London: DoH Publications, 2004.
22 Working group on 14-19 reform. *14-19 Curriculum and qualifications reform: final report of the working party on 14-19 reform [the Tomlinson report].* Nottingham: Department for Education and Skills Publications, 2004.
23 Lievens F, Coetsier P. Situational tests in student selection: an examination of predictive validity, adverse impact and construct validity. *Int J Sel Assess* 2002;10:245-57.
24 McManus IC, Vincent CA. Selecting and educating safer doctors. In: Vincent CA, Ennis M, Audley RJ, eds. *Medical accidents.* Oxford: Oxford University Press, 1993:80-105.
25 University of Cambridge Local Examinations Syndicate. FAQs on BMAT/TSA: "How do you know the tests are any good?" http://bmat.ucles-red.cam.ac.uk/background.html (accessed 25 Mar 2005).

(Accepted 6 June 2005)

Commentary: The benefits of aptitude testing for selecting medical students

Sandra Nicholson

Institute of Health Sciences Education, Queen Mary's School of Medicine and Dentistry, London E1 4NS
Sandra Nicholson *associate dean for admissions*
s.nicholson@qmul.ac.uk

The A level is the most common tool for assessing school leavers applying for higher education, including medicine. If medical school outcome is accurately predicted by A level grades, as described by McManus et al,[1] what place, if any, do aptitude tests have in the selection of medical students?

Applications for medical school from appropriately highly qualified candidates have increased year on year[2] until it has become increasingly difficult to discriminate between candidates with similar A level performance. Most medical schools wish to select future doctors using non-cognitive attributes alongside A levels, but procedures, such as interviewing, are time consuming and labour intensive. An urgent need is to reduce the number of candidates by initial screening that is appropriate, fair, and transparent but also gives added value to the process.

A further important reason for considering such testing initiatives is the concern that some groups are underrepresented in medicine because A level grades reflect educational background and social class.[3] Additional tests that can show intellectual ability or aptitude rather than achievement may be a valuable means to widen participation. These tests—some without undue reliance on a heavy science background—may be a useful adjunct to A levels where candidates offer a variety of

[As supplied by authors] *Selecting medical students*

Why intellectual aptitude tests probably aren't useful for selecting UK school-leaver entrants, and how A-levels could be

Supplementary Information

i). University achievement in relation to A-levels and aptitude.

a). The ISPIUA study of 1967.

b). The Westminster Study of 1975-1982.

c) The 1991 cohort study.

ii). Statistical problems in interpreting the relationship between A-level grades and university achievement.

iii) University performance of students in relation to aggregate school performance.

iv) Additional references from main text ('w' superscripts).

i) *University achievement in relation to A-levels and aptitude.*

a). *The ISPIUA study of 1967.* Initiated by the Committee of Vice-Chancellors and Principals in response to the proposed expansion of UK university education in the Robbins Report of 1963 [1], this large study looked at a random sample of 27,315 4th-term sixth-formers taking A-levels [2-4]. As well as collecting data on academic achievement and school assessments, the study also administered a three-hour Test of Academic Attitude (TAA), which separately assessed verbal ability and mathematical ability. 7,080 participants were known to have entered university in 1968, of whom 348 studied medicine. A further 2,315 participants entered university in 1969, with 118 studying medicine. The results for the 1968 entrants, shown in table 1, and the authors' conclusion, are clear: "TAA appears to add little predictive information to that already provided by GCE results [A-levels and O-lvels]and school assessment in general".

Table 1: **ISPIUA project (1967)**. Correlation of first-year degree performance (1968-69) and final degree performance (1971) with educational achievement and aptitude tests (1967). Note that medical students do not receive classed degrees, and have not finished by the third year of their course, and therefore are not included in the final column.

	First-year university performance (university tutor ratings)		Final degree
Correlation with:	Medicine (n=294)	All degree courses (n=5,985)	All degrees courses except medicine
Mean A-level grade	.35	.32	.36
Number of A-levels taken	.19	.14	.17
Number of O-levels taken	.20	.12	.16
School assessment	.23	.24	.26
Mathematical aptitude	.10	.01	.13
Verbal aptitude	.15	.01	-.02
Multiple correlation with all six predictors	*.41*	*.35*	*.42*

b). *The Westminster Study of 1975-1982.* All entrants to the clinical course at the Westminster Medical School between 1975 and 1982 were given the AH5 [5], a timed IQ test, as a measure of aptitude [6]. These students were followed up in 2002 as practising doctors, and a range of measures taken of professional attainment. Full details of correlations etc are available in the Supplementary Information of the paper (http://bmj.bmjjournals.com/cgi/content/full/327/7407/139/DC1) . A-level grades at the time of entry were predictive of professional achievement (time to Membership and Fellowship, achievement of Consultant posts, etc), although they did not predict research achievement or non-academic outcomes such as stress, burnout, and satisfaction with a medical career. The aptitude test provided little prediction of outcomes, and contributed no additional predictive power over and above that of A-level grades.

c) *The 1991 cohort study*. This study looked at 6901 applicants to UK medical schools in 1990, of whom 3333 were admitted to medical schools in 1991 or subsequently [7,8]. A sub-set of 786 applicants who were interviewed at three of the participating medical schools were given a series of timed psychometric tests, including an aptitude test consisting of an abbreviated version of the AH5 test [5,5]. Participants were followed-up at the end of the clinical course, as PRHOs [9],and in 2002 [10]. Outcomes for present purposes are performance in Basic Medical Science examinations at the end of year 2, finals performance (pass at expected time or delayed due to examination failure), and, in a subset of doctors who choose to take the examination, their first attempt at the first part of a UK postgraduate examination which is typically taken about 18 months after graduation.

Analyses of the aptitude test and A-levels in relation to performance on the course and afterwards have not previously been published, and are therefore provided here in somewhat more detail than for the previous two studies.

Table 2 summarises the correlations of each measure with mean A-level grade and overall AH5 performance. A-level grades were known for many more subjects than were the AH5 test results, and therefore to aid comparison, the correlations with A-level grades are provided both for all entrants and for all entrants who also took the AH5. Note that, as with the ISPIUA project, the verbal subtest of the aptitude test performs a little better than the non-verbal test.

Table 2: **1991 Cohort study (1990-2002)**. Correlations of outcome measures (Basic Medical Sciences course, Finals, and a postgraduate examination with mean A-level grade and performance on the aptitude test (abbreviated AH5, aAH5). Note that finals performance was only a binary measure, and there is relatively little power since most individuals qualified at their first sitting of the examination. Although a logistic regression is technically better for this binary measure, a correlation has been used for comparability with the other outcome measures.

	Basic Medical Sciences	Finals	Part 1 of postgraduate examination, Part 1, first attempt
Mean A-level grade (all entrants)	**.202** (p<.001, n=3112)	**.063** (p<.001, n=2510)	**.332** (p<.001, n=903)
Mean A-level grade (entrants taking aAH5)	**.210** (p<.001, n=756)	**.096** (p=.017, n=610)	**.371** (p<.001, n=239)
aAH5 (total)	**.044** (p=.227, n=766)	**.015** (p=.718, n=616)	**.123** (p=.057, n=240)
aAH5 (Verbal)	**.077** (p=.034, n=766)	**.045** (p=.262, n=616)	**.156** (p=.016, n=240)
aAH5 (Spatial)	**-.006** (p=.867, n=772)	**-.020** (p=.621, n=622)	**.048** (p=.463, n=240)

Supplementary figures 1, 2 and 3 (at the end of this **Supplementary Information)** show the relationship between A-level grades and overall aptitude test result with performance at Basic Medical Sciences, at Finals, and at the first attempt at the postgraduate examination.

We are unaware of any studies, other than the three described above, which have assessed both A-levels and aptitude tests as predictors of outcome at university, either in medicine or other subjects. We note in particular that the much cited 'pilot study' of aptitude tests in relation to A-level achievement which was carried out by the Sutton Trust [11,12]was relatively small, did not follow-up students to university, and concluded "the data provided no evidence that [the test] was able to assess potential for study at higher education, independently of a student's social and educational experiences".

iii) Statistical problems in interpreting the relationship between A-level grades and university achievement.

A major problem in interpreting the relationship between A-level grades, aptitude and university performance is what is known as 'restriction of range'. Selection often means that the range of achievement found in those selected is narrower than the range in those applying. The result is that any correlation will be attenuated from its true value for the population as a whole, with more variance due to measurement error and less due to systematic differences. To take an extreme, if one only looks at individuals gaining AAA at A-level then there can be no correlation with outcome, since there is no variance which can co-vary with outcome.

A variant argument is that since the correlation between aptitude tests and A-level grades is relatively low then they must be measuring different underlying cognitive components. That however is only the case if both are measured with high reliability and there is no restriction of range. In the Westminster study the correlation of A-level grade with aptitude was 0.285, a somewhat higher value than the 0.162 found in the 1991 Cohort Study in medical school applicants, but there was less restriction of range in the Westminster study. In the 1991 Cohort study the correlation was 0.238 in *interviewees*, not all of whom were accepted, so there was less restriction of range. In contrast in the ISPIUA project the correlation of A-levels and aptitude across *all* individuals (i.e. all of those taking A-levels) was a much higher 0.51 [13] (and that is not corrected for attenuation due to unreliability of the tests or additional restriction of range relative to the population as a whole, most of whom were not taking A-levels). It should also be remembered that very few social or psychological measures are perfectly reliable, and hence correlations of one, or even correlations above 0.5, are extremely rare, both due to the various causes of attenuation, and due to most social processes being inherently multifactorial. Correction for both measurement error and restriction of range is possible but is rarely carried out in practice, and many measures disparagingly described as 'weak', 'low' or 'poor' are actually quite good once they are corrected for such problems.

We should also add that we have recently seen data from two UK medical schools in which performance in examinations in the first three years was correlated with A-level grades, despite almost all having AAA, AAB or ABB grades; those with AAA had performed better than those with AAB who had in turn performed better than those with ABB. A-levels are still predictive of university outcome despite restriction of range.

To summarise, a low correlation between A-level grades and university outcome does not imply that A-levels are poor predictors of outcome and neither does a low correlation between A-levels and aptitude tests imply that separate cognitive processes are being measured. Both correlations suffer seriously from restriction of range, and attenuation due to unreliability, and cannot be taken at their face value.

iv. University performance of students in relation to aggregate school performance.

It has been hypothesised that the achievement of students who come from 'poorer' schools, where the aggregate level of performance at A-level is lower, will be higher at university than those with the same grades who come from schools which on average are performing at a higher level. The assumption is that academic achievement in adversity, where few students achieve well at A-level, perhaps because of lower quality teaching, will be a better predictor of success than similar achievement as a result of high quality teaching. Although a reasonable argument, it does not take into account the possibility that schools which are poorer achievers at A-level are so because the students in them are of lower intellectual ability on average. In general the hypothesis is usually put forward without any data in its support. However the recent analyses of HEFCE suggest that the hypothesis is without empirical foundation.

HEFCE's analysis of 79,005 18-year-olds entering university in 1997-1998 and followed until 2000-2001 (http://www.hefce.ac.uk/pubs/hefce/2003/03_32.htm) divided schools into four quartiles, from those with the highest aggregate A-level results to those with the lowest aggregate A-level results. Supplementary figure 4 shows the outcome at university, in terms of proportion of II.i or I degrees, in relation to individual A-level performance, and in relation to aggregate school performance. There is a small difference between those in the highest quartile and the other three quartiles, but no obvious difference between the three lowest quartiles. Further analysis of the data by HEFCE found that quartile of school achievement was strongly confounded with school type, all of the Independent schools being in the highest quartile of school achievement. When school type was taken into account, there was no effect of average level of school achievement. Therefore it has to be concluded that aggregate school achievement is not a predictor of university outcome in individual pupils, despite the seductive attractiveness of the original hypothesis.

Reference List (Web material only)

1. Committee on Higher Education. Higher Education: Report of the Committee appointed by the Prime Minister under the chairmanship of Lord Robbins 1961-63, Cmnd 2154. London: HMSO, 1963.

2. Choppin BHL, Orr L, Kurle SDM, Fara P, James G. The prediction of academic success. Slough: NFER Publishing Company Ltd, 1973.

3. Choppin BHL, Orr L, Fara P, Kurie SDM, Fogelman KR, James G. After A-Level? A study of the transition from school to higher education. Slough: NFER Publishing Company Ltd, 1972.

4. Choppin B. Admission tests for admission to university: the British experience. In Mitter W, ed. *The use of tests and interviews for admission to higher education*, Windsor: National Foundation for Educational Research, 1979.

5. Heim AW. AH5 group test of high-grade intelligence. Windsor: NFER, 1968.

6. McManus IC, Smithers E, Partridge P, Keeling A, Fleming PR. A levels and intelligence as predictors of medical careers in UK doctors: 20 year prospective study. *British Medical Journal* 2003;**327**:139-42.

7. McManus IC, Richards P, Winder BC, Sproston KA, Styles V. Medical school applicants from ethnic minorities: identifying if and when they are disadvantaged. *British Medical Journal* 1995;**310**:496-500.

8. McManus IC, Richards P, Winder BC. Intercalated degrees, learning styles, and career preferences: prospective longitudinal study of UK medical students. *British Medical Journal* 1999;**319**:542-6.

9. McManus IC, Winder BC, Paice E. How consultants, hospitals, trusts and deaneries affect pre-registration house officer posts: a multilevel model. *Medical Education* 2002;**36**:35-44.

10. McManus IC, Keeling A, Paice E. Stress, burnout and doctors' attitudes to work are determined by personality and learning style: A twelve year longitudinal study of UK medical graduates. *BMC Medicine* 2004;**2:29**.

11. McDonald AS, Newton PE, Whitton C, Higgs SE. A pilot of aptitude testing for university entrance. London: National Foundation for Educational Research, 2000.

12. McDonald AS, Newton PE, Whetton C, Benefield P. Aptitude testing for university entrance: A literature review. London: National Foundation for Educational Research, 2000.

13. Whetton C, McDonald A, Newton P. Aptitude testing for university entrance. Slough: National Foundation for Education Research **www.nfer.ac.uk/research/papers/PaperforRio090301.doc**, 2001.

iv) Additional references from main text. (‘w’ superscripts).

w1 Anderson J, Hughes D, Wakeford RE. Medical student selection: a tentative attempt to establish a code of practice. British Medical Journal 1980;i:1216-8.

w2 McManus IC. Factors affecting likelihood of applicants being offered a place in medical schools in the United Kingdom in 1996 and 1997: retrospective study. British Medical Journal 1998;317:1111-6.

w3 UCLES (University of Cambridge Local Examinations Syndicate): Thinking Skills Assessment website. http://tsa.ucles.org.uk/index.html (Accessed 25 March 2005)

w4 McDonald AS, Newton PE, Whetton C, Benefield P. Aptitude testing for university entrance: A literature review. London: National Foundation for Educational Research, 2000.

w5 Bagg DG. A-levels and university performance. Nature 1970;225:1105-8.

w6 Sear K. The correlation between A-level grades and degree results in England and Wales. Higher Education 1983;12:609.

w7 Buxton RSJ. Medical students leaving before qualification. British Journal of Medical Education 1993;7:155-6.

w8 Green A, Peters TJ, Webster DJT. Student selection: are the school-leaving A-level grades in biology and chemistry important? Medical Education 1993;27:22-5.

w9 Peers IS,.Johnston M. Influence of learning context on the relationship between A- level attainment and final degree performance: a meta-analytic review. British Journal of Educational Psychology 1994;64:1-17.

w10 Simpson KH,.Budd K. Medical student attrition: a 10 year survey in one medical school. Medical Education 1996;30:172-8.

w11 Buxton RSJ. Medical students leaving before qualification. British Journal of Medical Education 1973;7:155-6.

w12 Bowler I. Ethnic profile of the doctors in the United Kingdom. British Medical Journal 2004; 329:583-4.

w13 Neame RLB, Powis DA, Bristow T. Should medical students be selected only from recent school-leavers who have studied science? Medical Education 1992;26:433-40.

w14 Frey MC,.Detterman DK. Scholastic assessment or g? The relationship between the Scholastic Assessment Test and general cognitive ability. Psychological Science 2004;15:373-8.

w15 Moutafi J, Furnham A, Crump J. Demographic and personality predictors of intelligence: a study unsing the NEO personality inventory and the Myers-Briggs Type Indicators. European Journal of Personality 2003;17:79-94.

w16 Balin S, Case R, Coombs JR, Daniels LB. Common misconceptions of critical thinking. Journal of Curriculum Studies 1999;31:269-283.

w17 Balin S, Case R, Coombs JR, Daniels LB. Conceptualising critical thinking. Journal of Curriculum Studies 1999;31:285-302.

w18 Heim AW. AH5 group test of high-grade intelligence. Windsor: NFER, 1968.

w19 McManus IC, Richards P, Winder BC, Sproston KA, Styles V. Medical school applicants from ethnic minorities: identifying if and when they are disadvantaged. British Medical Journal 1995;310:496-500.

w20 McManus IC, Richards P, Winder BC. Intercalated degrees, learning styles, and career preferences: prospective longitudinal study of UK medical students. British Medical Journal 1999;319:542-6.

w21 McManus IC, Winder BC, Paice E. How consultants, hospitals, trusts and deaneries affect pre-registration house officer posts: a multilevel model. Medical Education 2002;36:35-44.

w22 McManus IC, Keeling A, Paice E. Stress, burnout and doctors' attitudes to work are determined by personality and learning style: A twelve year longitudinal study of UK medical graduates. BMC Medicine 2004;2:29.

w23 Ferguson E, Sanders A, O'Hehir F, James D. Predictive validity of personal statements and the role of the five factor model of personality in relation to medical training. Journal of Occupational and Organizational Psychology 2000:73:321-344.

w24 McManus IC. Social class data are problematic to interpret [letter]. British Medical Journal 2004;329:800-1.

w25 Seyan K, Greenhalgh T, Dorling D. The standardised admission ratio for measuring widening participation in medical schools: analysis of UK medical school admissions by ethnicity, socioeconomic status, and sex. British Medical Journal 2004;328:1545-6.

w26 Greenhalgh T, Seyan K, Boynton P. "Not a university type": focus group study of social class, ethnic, and sex differences in school pupils' perceptions about medical school. British Medical Journal 2004;328:1541-4.

w27 McManus IC. The social class of medical students. Medical Education 1982;16:72-5.

w28 Patterson, F, Ferguson E, Norfolk T, Lane P. A new selection system to recruit general practice registrars: preliminary findings from a validation study. British Medical Journal 2005;330:711-714

w29 Lowe M, Kerridge I, Bore M, Munro D, Powis D. Is it possible to assess the "ethics" of medical school applicants? J Med Ethics 2001;27:404-8.

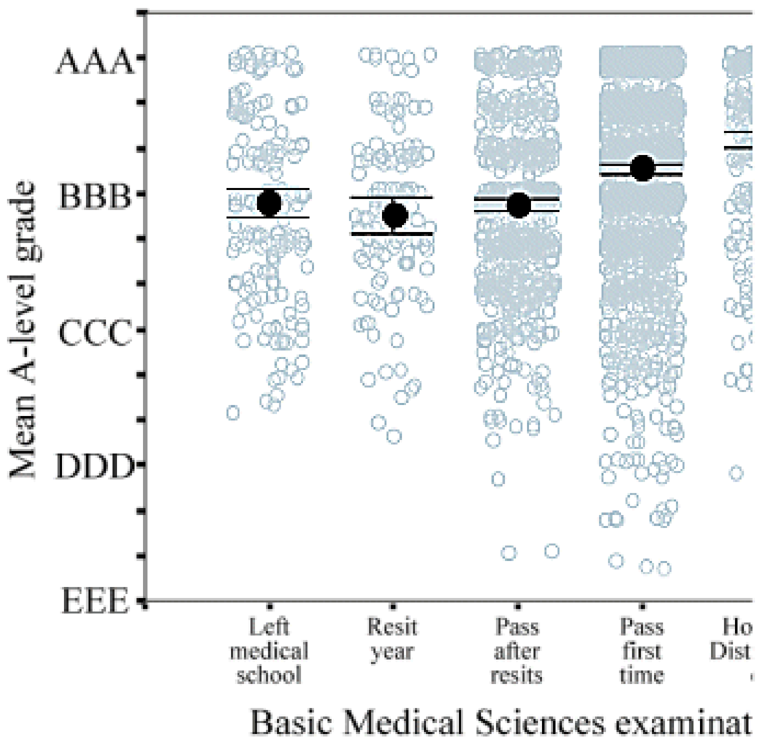
w30 Albanese MA, Snow MH, Skochelak SE, Hugget KN, Farrell PM. Assessing personal qualities in medical school admissions. Academic Medicine 2003;78:313-21.

w31 Association of American Medical Colleges and National Board of Medical Examiners. Embedding professionalism in medical education: Assessment as a tool for implementation. Philadelphia: National Board of Medical Examiners, 2003.

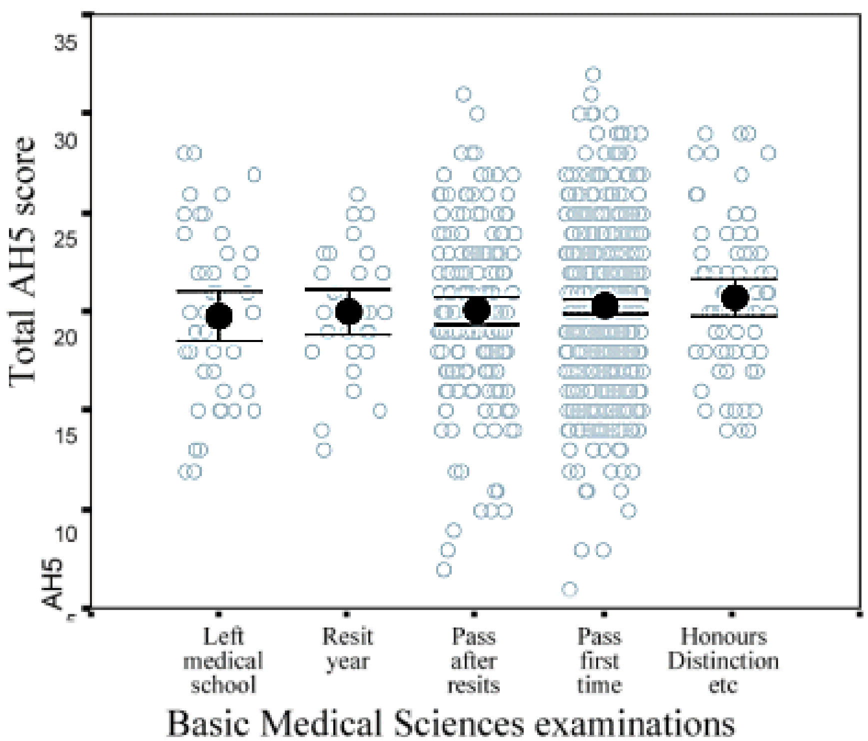
w32 McManus IC. From selection to qualification: How and why medical students change. In: Allen I, Brown P, Hughes P, eds. Choosing tomorrow's doctors, pp 60-79. Policy Studies Institute, London, 1997.

Supplementary figure 1

a).

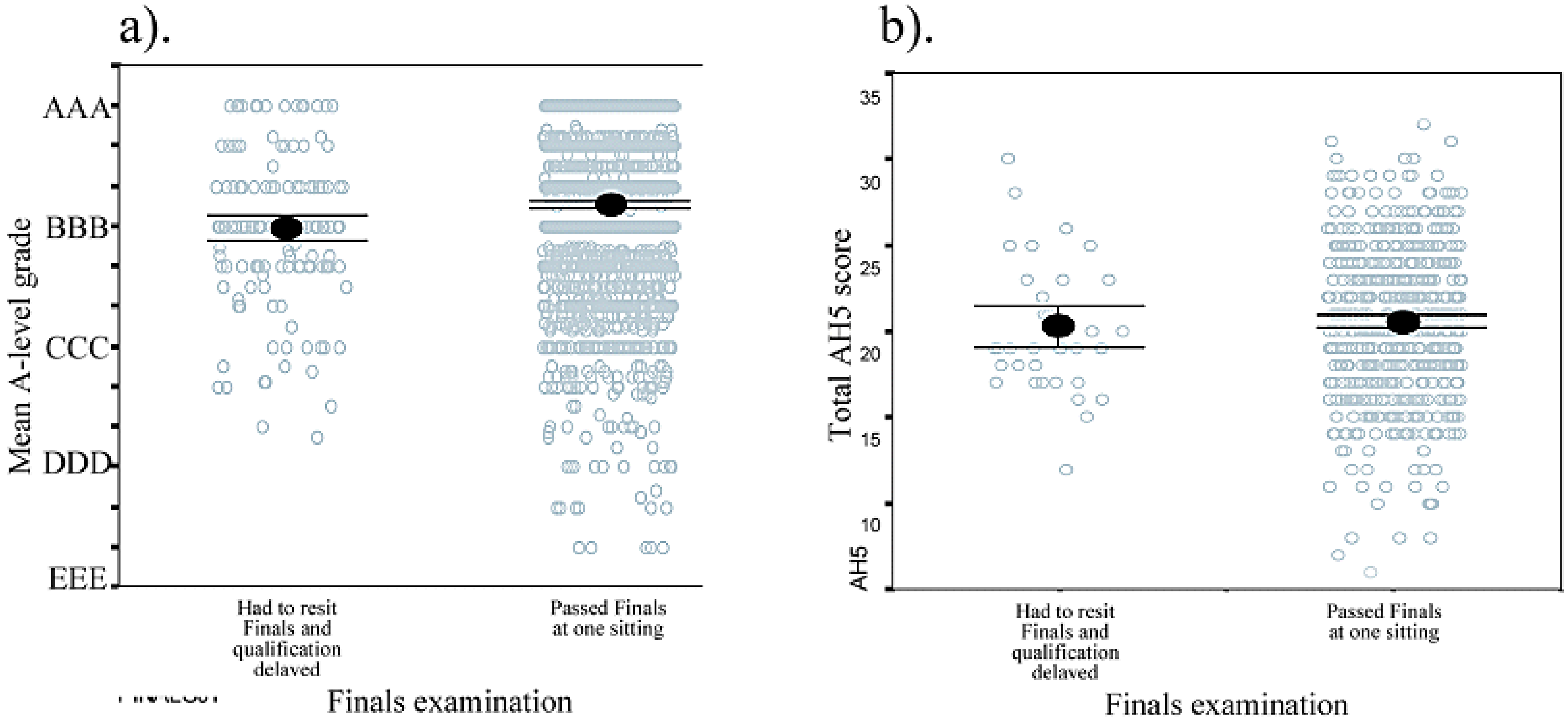


b).



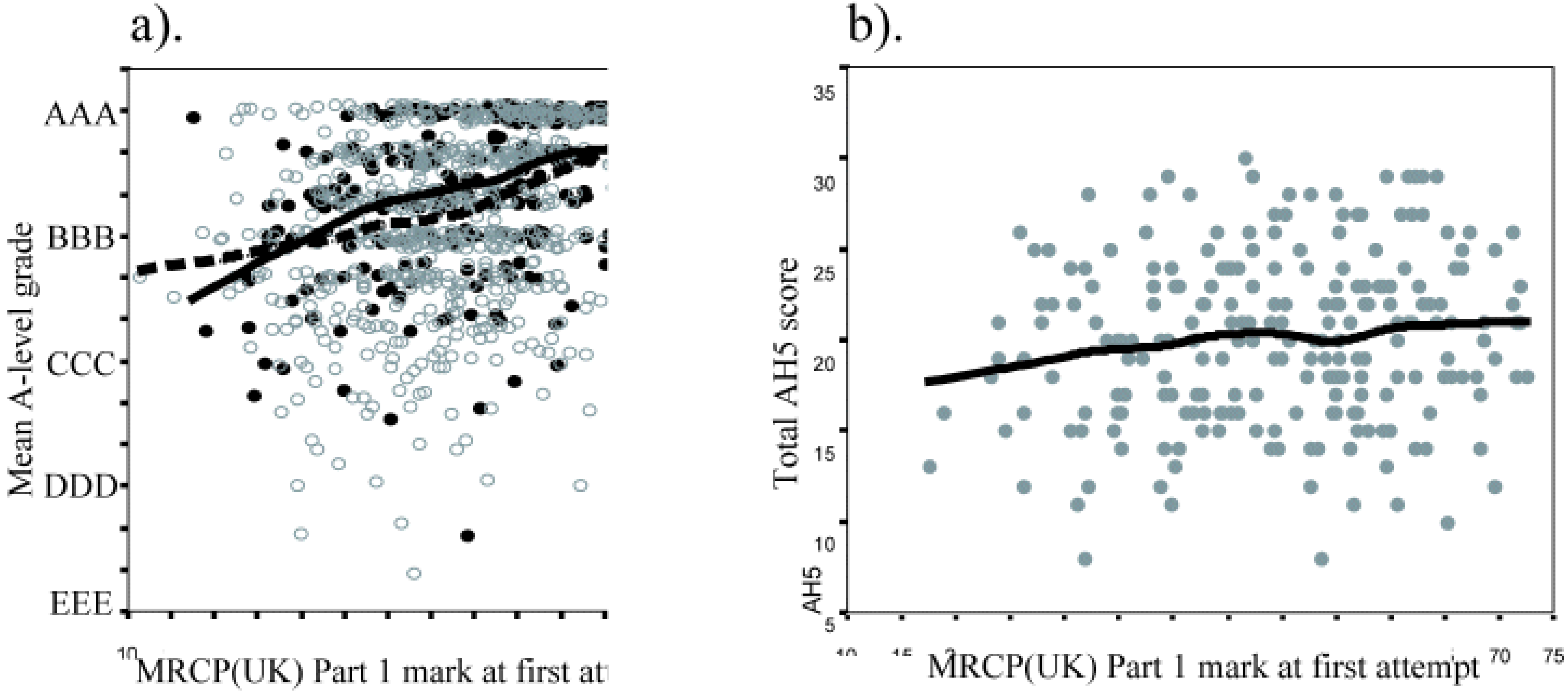
Supplementary Figure 1. The abscissa of both figures shows performance of students in the 1991 cohort study on Basic Medical Sciences (BMS) examinations, typically reflecting performance in the first two years of the medical curriculum. BMS results are only reported in terms of five simple categories because of differences in measurement in different medical schools. The left-hand figure (a.) shows the relationship to mean A-level grade, individual subjects being shown as pale grey circles, and means (± 1 standard error) superimposed in black. The right-hand figure (b.) shows the relationship to the overall AH5 score, with pale grey circles for individual subjects, and means (± 1 standard error) superimposed in black. Individual points have been given a slight random jitter to right or left in order that points do not superimpose upon one another.

Supplementary figure 2



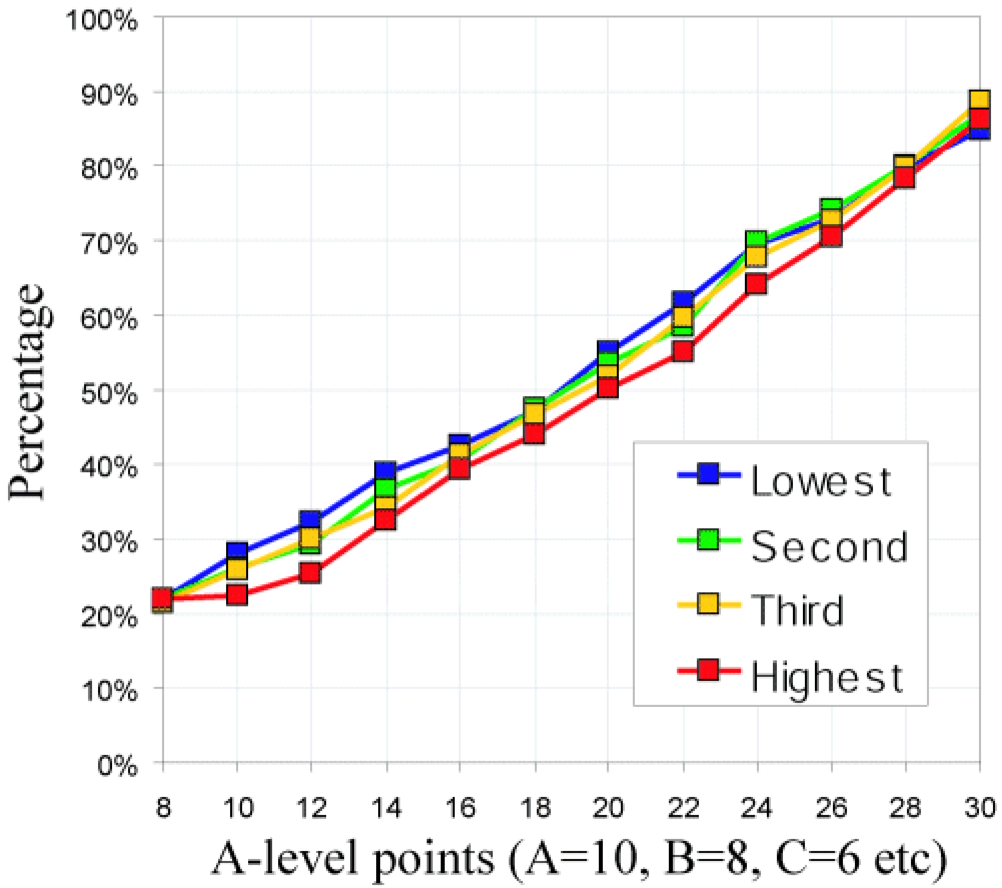
Supplementary Figure 2. The abscissa of both figures shows performance of students in the 1991 cohort study at their final examinations. Because of differences in method of examining between medical schools, the results are merely expressed as whether candidates passed their finals at the first sitting, and hence qualified, or instead had to resit finals six months or so later, and hence qualified late. The left-hand figure (a.) shows the relationship to mean A-level grade, individual subjects being shown as pale grey circles, and means (± 1 standard error) superimposed in black. The right-hand figure (b.) shows the relationship to the overall AH5 score, with pale grey circles for individual subjects, and means (± 1 standard error) superimposed in black. Individual points have been given a slight random jitter to right or left in order that points do not superimpose upon one another.

Supplementary figure 3



Supplementary Figure 3. The abscissa of both figures shows performance of those students in the 1991 cohort study who had taken the 1st Part of the postgraduate examination on their first attempt at the examination. The left-hand figure (a.) shows the relationship to mean A-level grade, individual subjects being shown as circles. The superimposed lines are lowess curves. The solid black circles and the solid line are for those subjects for whom AH5 results were available, whereas the open grey circles and dashed line are for subjects who had not taken the AH5 test. Individual points have been given a slight vertical jitter to prevent subjects with exactly the same A-level grades being superimposed upon one another.

Supplement



Supplementary Figure 4. The proportion of university entrants with particular A-level grades who gained a II.i or I degree, in relation to the overall performance of their secondary school at A-level. School performance is ordered from the highest quartile (schools with the highest average grades), through to the lowest quartile (schools with the lowest average grades). Redrawn from the HEFCE data reported at www.hefce.ac.uk/pubs/hefce/2003/03_32.htm.