GL Education®

TECHNICAL REPORT – INTERNATIONAL EDITION

CAT4®

COGNITIVE ABILITIES TEST

Contents

CAT4 UK EDITION

Test reliability

The reliability of a test is a measure of the consistency of a student's test scores over repeated testing, assuming conditions remain the same – that is, there was no fatigue, learning effect or lack of motivation. Tests with poor reliability might result in very different scores for a student across two test administrations.

The reliability of the test was estimated using the Cronbach's Alpha formula which produces values ranging from 0 to 1. Values above 0.80 are considered to be very good. The reliability values for the various *CAT4* batteries are given in the table below, and all show that the tests are very reliable. These are based on students who took part in the UK standardisation.

CAT4 level	*CAT4* reliability				
	Verbal Reasoning Battery	Quantitative Reasoning Battery	Nonverbal Reasoning Battery	Spatial Ability Battery	Overall *CAT4*
Level X	0.93	0.91	0.87	0.83	0.95
Pre-A	0.82	0.81	0.78	0.67	0.90
A	0.91	0.91	0.90	0.87	0.97
B	0.89	0.90	0.90	0.88	0.96
C	0.86	0.91	0.87	0.85	0.96
D	0.90	0.91	0.89	0.86	0.96
E	0.89	0.88	0.86	0.88	0.96
F	0.89	0.87	0.85	0.88	0.96
G	0.90	0.84	0.85	0.86	0.95
Average	**0.89**	**0.88**	**0.87**	**0.84**	**0.95**

For interpreting the score of an individual student, the standard error of measurement (*SEM*) is a more useful statistic than a reliability coefficient. It indicates how large, on average, the fluctuations in standard scores may be. The *SEM* for the Verbal Reasoning Battery is 5.0, which indicates that there is a 68% chance that the student's true verbal SAS will be in the range +/- 5.0. For example, for an average-performing student with a verbal SAS of 100, there is a 68% chance that his or her true verbal score is in a range from 95 to 105.

CAT4 level	*CAT4* Standard error of measurement (*SEM*)				
	Verbal Reasoning Battery	Quantitative Reasoning Battery	Nonverbal Reasoning Battery	Spatial Ability Battery	Overall *CAT4*
Average	5 .0	5.2	5.4	6.0	3.4

However, most tests show the 90% chance or confidence bands. For values around the average, the 90% confidence band is as follows:

CAT4 level	CAT4 90% confidence band				
	Verbal Reasoning Battery	Quantitative Reasoning Battery	Nonverbal Reasoning Battery	Spatial Ability Battery	Overall CAT4
Average	+/- 8	+/- 9	+/- 9	+/- 10	+/- 6

For example, for an average-performing student with a verbal SAS of 100, there is a 90% chance that the true verbal score is in a range from 92 to 108.

Test re-test reliability

A study of 3,883 students who took Level D and subsequently took Level F two years later showed the correlation for the mean *CAT4* SAS between the two time points was high at 0.88. The correlations for the overall mean *CAT4* SAS and the four batteries are shown in the table below:

	Correlation between Level D and Level F SAS
Mean *CAT4*	0.88
Verbal	0.80
Quantitative	0.75
Nonverbal	0.71
Spatial	0.74

The results showed a high level of consistency and: 62% of students had mean *CAT4* scores within +/- 5 SAS points; 90% of students had mean *CAT4* scores within +/- 10 SAS points.

Cognitive Abilities Test and National Test indicators

There has always been a significant and positive correlation between a student's scores in reasoning tests and their school performance, as measured by national tests or public examinations. The link may be assumed to exist because much school activity is concerned with the application of reasoning abilities in the initial learning of curriculum content, and then building on and recombining existing knowledge as learning progresses.

The indicators that feature in reports for the Cognitive Abilities Test are derived by tracking the progress of large and representative samples of students over time. Through this process, we can determine the actual relationship between *CAT4* scores and students' subsequent attainment in national tests and examinations.

Through statistical analysis of the matched datasets, we are able to provide indicated or typical outcomes for each student based on the students' *CAT4* scores. These indicators can also be aggregated to provide indicated outcomes for the cohort and school or college as a whole. These indicators are updated regularly to keep them in line with national trends of performance in national tests and examinations.

Key Stage 2 National Test indicators: England

The KS2 indicators are derived from an analysis of the relationship between *CAT4* scores from Level A to Level C and KS2 test results at age 11 from a large and nationally representative sample of around 24,000 students taking the KS2 SATS in 2019. This relationship between CAT4 scores and KS2 SATS is also used to estimate the retrospective KS2 indicators. These indicators are updated regularly as we get new data.

Correlations of *CAT4* and KS2 scaled scores

There is a strong relationship between *CAT4* scores and Key Stage 2 outcomes. The strength of the relationship between two variables can be measured by a statistic called the correlation coefficient. A value of zero indicates no relationship between the two measures, whereas a value of one indicates a perfect positive relationship. The table below shows the correlation coefficients between *CAT4* standard age scores (SAS) and students' subsequent KS2 scaled score outcomes.

KS2 SATS scaled scores	Mean *CAT4* score	Verbal SAS	Quantitative SAS	Nonverbal SAS	Spatial SAS
Mathematics	0.74	0.63	0.71	0.62	0.59
Reading	0.66	0.68	0.57	0.54	0.49
Grammar, Punctuation and Spelling	0.69	0.69	0.64	0.56	0.50

The correlations are all highly significant. The Mathematics outcomes tend to have their highest correlation with the mean *CAT4* SAS. The *CAT4* Verbal Reasoning score gives a slightly higher or similar correlation than the mean *CAT4* score for English Reading, and Grammar, Punctuation and Spelling.

The graph below illustrates the relationship between the mean *CAT4* score and the KS2 Mathematics scaled scores. It shows the most likely scaled score and the score if the student is challenged. We can see that the scaled scores increase as the *CAT4* scores increase.

CAT4 scores and Maths scaled scores
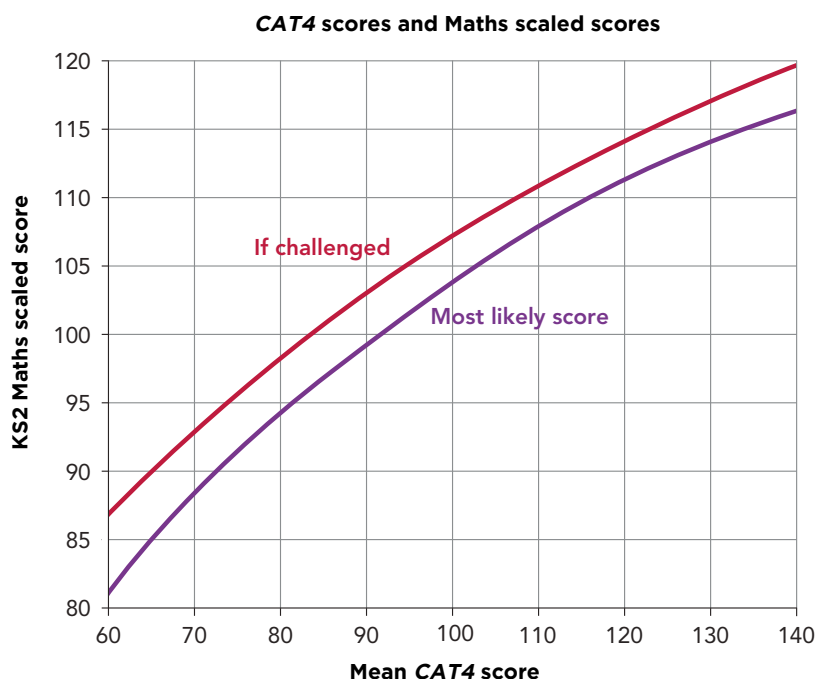
For example, a student with a mean *CAT4* score of 90, the 'most likely' Mathematics scaled score is 99 and the 'if challenged' threshold is 103. Not all students with a mean *CAT4* score of 90 will get a Mathematics scaled score of 99. The 'most likely' score is an average, so around half of the students with mean *CAT4* scores of 90 will obtain a Mathematics scaled score below 99; 25% of the students will obtain a Mathematics scaled score of between 99 and 102; and 25% of the students will obtain an 'if challenged' score of 103 or above.

Likelihood of Key Stage 2 indicated standard

The graph below illustrates the proportion of students achieving a scaled score of 100 (the government's expected standard) or the high score of 110 for Mathematics for each mean *CAT4* score. We can see that the higher the mean *CAT4* score, the greater the proportion of students who achieve the government's benchmark or above. For example, 58% of students with a mean *CAT4* score of 90 obtained the expected standard of 100 or above in Mathematics; in contrast, about 95% of students with a mean *CAT4* score of 110 achieved this.
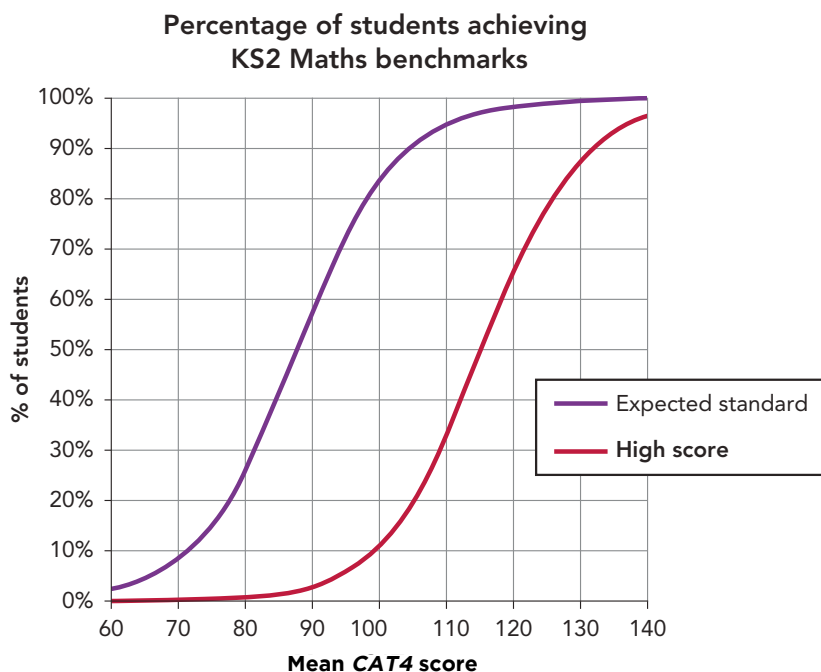
**Percentage of students achieving
KS2 Maths benchmarks**

The chart below illustrates the relationship between the Verbal *CAT4* score and the KS2 English Reading benchmarks.

**Percentage of students achieving
KS2 English Reading benchmarks**
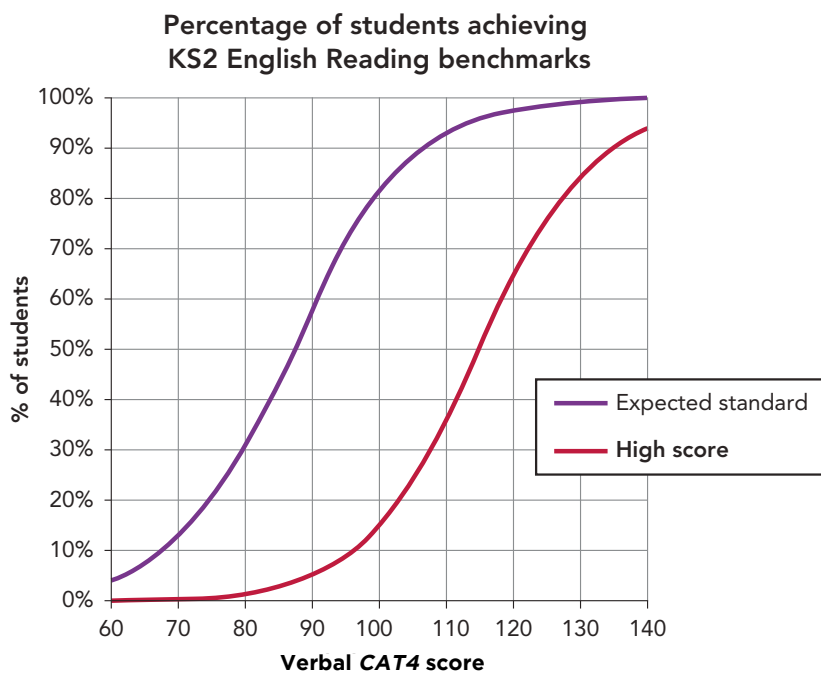
The chart below illustrates the relationship between the Verbal *CAT4* score and the KS2 English Spelling, and Grammar (*SPAG*) benchmarks.

**Percentage of students achieving
KS2 English SPAG benchmarks**
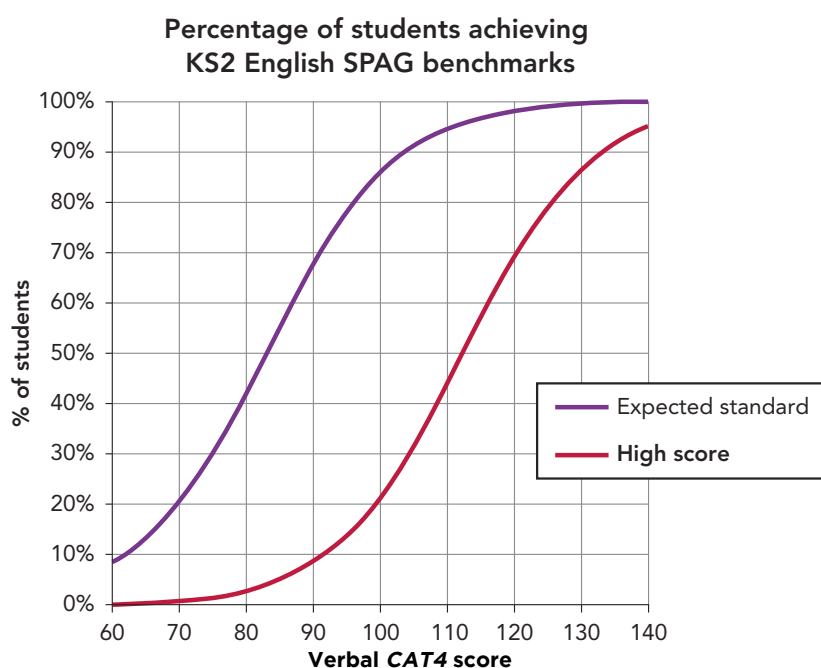
KS2 indicators for groups of students

The table below illustrates how the group/class indicators have been calculated for a fictitious group of five students and shows the probability of obtaining different KS2 Mathematics benchmarks.

	Mean *CAT4* score	Most likely scaled score achieved in Mathematics	Probability of students reaching	
			Expected standard = 100	High score = 110
Student 1	85	97	41%	2%
Student 2	95	102	73%	6%
Student 3	106	106	92%	23%
Student 4	109	108	95%	31%
Student 5	111	108	96%	37%
Average			80%	20%
Number of students achieving:			4	1

The individual student indicators do not show any of these five students likely to obtain a high scaled score benchmark of 110 or more. However, some students have a high chance of achieving this, e.g. student 5 has a 37% chance of obtaining a high score of 110 or more. Overall for this group of five students we expect 20% (i.e. one out of the five students) to achieve the high score. As an illustration, if your group has 10 students all with mean *CAT4* scores of 106, the most likely outcome for each of these 10 students individually is a scaled score of 106. However, it is likely that 23% of these students (i.e. two out of the 10 students) will achieve the high score.

The group level indicators are the average of the probabilities for all students in the group. Our research has shown that this method provides the most accurate set of group level indicators. However, group indicators are extremely sensitive to variations in the number of students in the group, and may be very unstable for groups of less than 30 students. Group indicators should only ever be taken as a rough guide to the possible future performance of a class.

Key Stage 2 National Test indicators: Wales

The *CAT4* KS2 reports for Wales show estimates of the Literacy and Numeracy National Tests age-standardised scores as well as estimates of teacher assessment levels.

The table below shows the correlations between *CAT4* and the Year 6 National Tests and teacher assessments. This is based on a study of around 2,500 students who completed *CAT4* and the National Tests in Wales.

Welsh test	Mean *CAT4* score	Verbal SAS
Literacy	0.68	0.70
Numeracy (Procedural)	0.70	0.60
Numeracy (Reasoning)	0.64	0.54
English teacher assessment	0.66	0.67
Maths teacher assessment	0.69	0.62
Science teacher assessment	0.65	0.63
Welsh 2nd subject teacher assessment	0.54	0.55

The correlations are all highly significant. The Mathematics and Science outcomes tend to have their highest correlation with the mean *CAT4* SAS. The *CAT4* Verbal Reasoning score alone gives a slightly higher correlation than the mean *CAT4* score for Literacy, English and Welsh 2nd subject.

GCSE indicators

The GCSE indicators are derived from an analysis of the relationship between *CAT4* scores from Level D and above and GCSE examination results at age 16 for a large and nationally representative sample of around 91,000 students in 2019. These indicators are updated regularly as we get new data.

Correlations of *CAT4* and GCSE grades

As already stated, the strength of the relationship between two variables can be measured by a statistic called the correlation coefficient. A value of zero indicates no relationship between the two measures, whereas a value of one indicates a perfect positive relationship. The table below shows the correlation coefficients between *CAT4* standard age scores and pupils' subsequent GCSE outcomes.

	Mean *CAT4* score	Verbal SAS	Quantitative SAS	Nonverbal SAS	Spatial SAS
Attainment 8*	0.72	0.67	0.64	0.61	0.57
Art and Design	0.48	0.44	0.38	0.41	0.42
Biology	0.62	0.57	0.53	0.49	0.47
Business Studies	0.56	0.45	0.52	0.49	0.40
Chemistry	0.57	0.50	0.50	0.45	0.43
Citizenship	0.51	0.52	0.45	0.41	0.35
Computer Studies	0.65	0.60	0.56	0.53	0.51
Design and Technology	0.55	0.47	0.51	0.45	0.46
Drama	0.55	0.55	0.45	0.47	0.42
English Language	0.62	0.62	0.53	0.51	0.46
English Literature	0.58	0.57	0.50	0.48	0.43
Food and Nutrition	0.61	0.59	0.53	0.51	0.47
French	0.53	0.54	0.45	0.43	0.38
Geography	0.68	0.65	0.59	0.56	0.52
German	0.53	0.54	0.45	0.42	0.38
History	0.60	0.59	0.52	0.48	0.43
ICT	0.52	0.43	0.49	0.46	0.38
Maths	0.78	0.66	0.72	0.66	0.63
Media Studies	0.50	0.41	0.48	0.42	0.38
Music	0.56	0.55	0.50	0.44	0.45
Physical Education	0.60	0.56	0.52	0.49	0.46
Physics	0.60	0.52	0.52	0.47	0.46
Religious Education	0.53	0.52	0.46	0.44	0.37
Science Combined	0.66	0.59	0.56	0.55	0.50
Sociology	0.48	0.39	0.48	0.40	0.34
Spanish	0.45	0.44	0.38	0.37	0.35
Statistics	0.72	0.60	0.60	0.67	0.57

Attainment 8 score is a summary score used in England.

The correlations are all highly significant. Most GCSE outcomes tend to have their highest correlation with mean *CAT4* score. The exceptions are English Language and English Literature where the *CAT4* Verbal Reasoning score gives a slightly higher correlation than mean *CAT4* score.

Likelihood of GCSE indicated grades

The example below illustrates the probabilities of achieving the various GCSE 9-1 grades in Mathematics (U is ungraded) for a student with a mean *CAT4* score of 100. The indicators are not precise: they indicate the outcomes expected for students with a particular *CAT4* score making average progress in a typical secondary school.

The 'most likely grade achieved' is reported to one decimal place. In this case the student is expected to be on the top end of grade 4 as he has a 52% chance of achieving grade 4 or below and a 48% chance of achieving grade 5 or above, so the expectation is that the student is near the grade 4/5 boundary.

Student name	Mean *CAT4* score	Mathematics GCSE grades – probabilities									Most likely grade achieved
		U/1	2	3	4	5	6	7	8	9	
John Sims	100	2%	4%	11%	34%	30%	11%	5%	2%	0%	4.9

The example below illustrates the probabilities of achieving the various GCSE A*-G grades in History (U is ungraded) for a student with a mean *CAT4* score of 100.

The 'most likely grade achieved' is grade C with the student having a 64% chance of achieving grade C or below and a 34% chance of achieving grade B or above.

Student name	Mean *CAT4* score	History GCSE grades – probabilities most likely									Most likely grade achieved
		U	G	F	E	D	C	B	A	A*	
John Sims	100	2%	2%	5%	11%	18%	26%	22%	11%	3%	C

GCSE grade indicators for groups of students

The table below illustrates how the group/class indicators have been calculated for a fictitious class with five students and shows the most likely grade achieved and the probabilities associated with getting different Mathematics 9-1 grades. The group indicator is an average of the individual student outcomes and probabilities. A similar method is used for subjects using the A*-G grades.

Using individual student grade estimates to provide information about the overall class or group grade outcomes will in most cases lead to underestimating the number of students likely to get both the higher and lower GCSE grades.

Student	Mean CAT4 score	Attain- ment 8 score	Mathematics GCSE grades – probabilities									Most likely grade achieved
			U/1	2	3	4	5	6	7	8	9	
1	70	11	78%	14%	5%	2%	0%	0%	0%	0%	0%	1
2	85	29	21%	26%	27%	20%	5%	1%	0%	0%	0%	2.8
3	100	46	2%	4%	11%	34%	30%	11%	5%	2%	0%	4.9
4	115	63	0%	0%	1%	6%	17%	23%	28%	18%	6%	6.9
5	140	79	0%	0%	0%	0%	0%	1%	3%	13%	83%	9
Group indicator (average)		46	20%	9%	9%	12%	11%	7%	7%	7%	18%	4.9

The group level indicators are the average of the probabilities for all students in the group. Our research has shown that this method provides the most accurate set of group level indicators. However, group indicators are extremely sensitive to variations in the number of students in the group, and may be very unstable for groups of less than 30 students. Group indicators should only ever be taken as a rough guide to the possible future performance of a class.

CAT4 and GCSE Attainment 8

The graph below illustrates the relationship between CAT4 score and the Attainment 8 score.

CAT4 score and Attainment 8
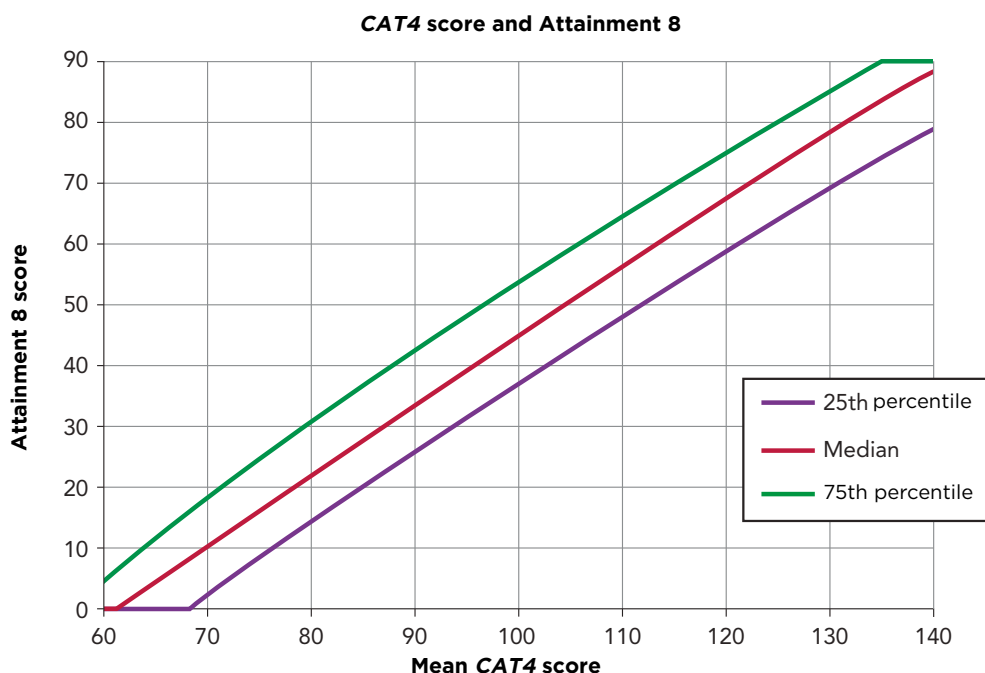
For example, for a student with a mean CAT4 score of 90, the most likely Attainment 8 is 42 and the 'if challenged' score is 49. Not all students with a mean CAT4 score of 90 will get an Attainment 8 score of 35.

Around half the students will get an Attainment 8 score below 35, with around 25% of the students obtaining an Attainment 8 score of less than 26 – the bottom 25th percentile. Around 25% of students will obtain the 'if challenged' score of 43 and above.

CAT4 and GCSE grades A*-G

Wales is retaining the current A*-G grading system; but in Northern Ireland the GCSE grading system is currently the same as for England, using the mixture of A*-G and 9-1 grades. A new structure based on a revised A*-G grading system was implemented in Northern Ireland in summer 2019. The new A* aligns closely to grade 9, and a new C* grade is equivalent to grade 5.

The graph below illustrates the proportion of students achieving five+ GCSE grades 9-4 (A*-C) including English and Mathematics for each mean CAT4 score. We can see that the higher the mean CAT4 score, the greater the proportion of students who achieve five or more 9-4 (A*-C) grades. For example, only 17% of students with a mean CAT4 score of 85 obtain five+ A*-C grades; in contrast, about 89% of students with a mean CAT4 score of 115 achieve five+ 9-4 (A*-C) grades.

**Probability of five or more GCSEs at grades 9-4 (A*-C)
including English and Mathematics**
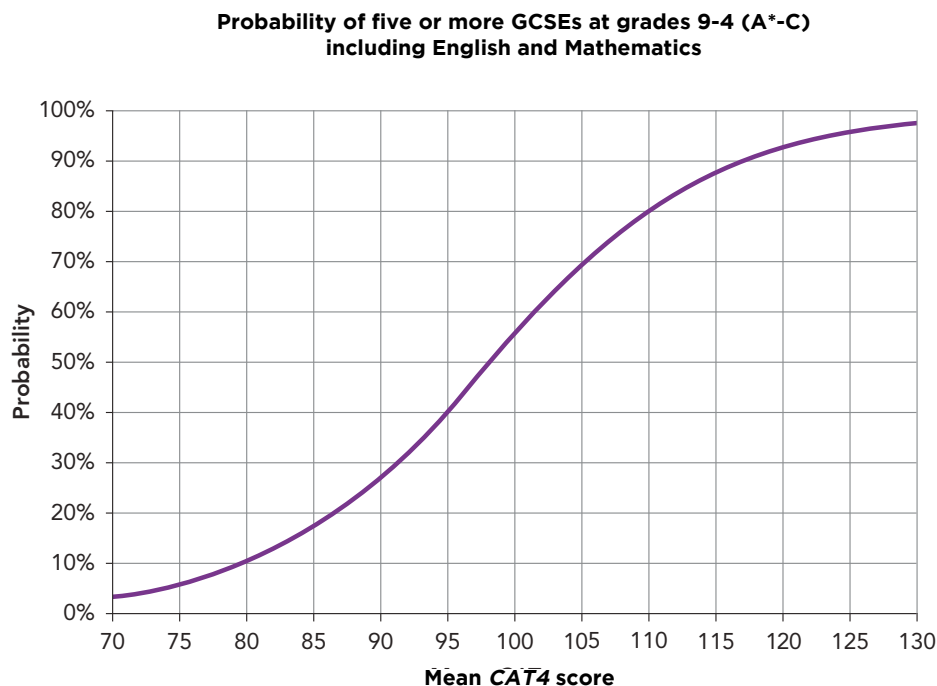
Setting targets

The above confirms the need for suitably cautious interpretation when using the indicators with staff and parents, and particularly if sharing them with individual students. In the latter context, we would advise that school staff follow the established best practice of schools, using the results for mentoring and target-setting purposes by:

* stressing to students that the indicators are a statistical prediction, not a prophecy of their actual Key Stage or GCSE results;

* emphasising to students the range of outcomes that could be achieved;

* emphasising the importance of the students' motivation and effort in determining the grade they obtain, identifying any areas in which the student requires greater support from the teacher;

* not using the indicators to label students as actual or potential 'failures';

* setting the indicators in the context of all other known relevant factors and other assessment information, thus making sure targets are reasonable.

International Baccalaureate (IB) indicators

IB grade exam results were collected from 744 students who had done the CAT4 test. Results were collected for:

G1- Studies in Language and Literature

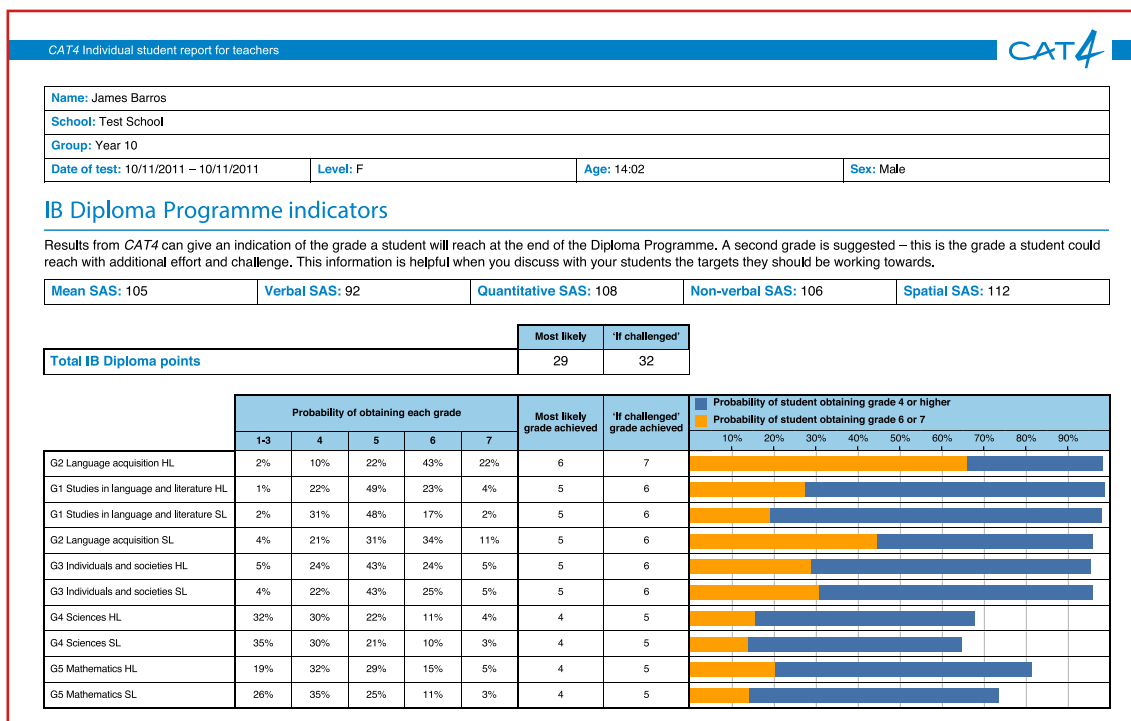G2- Language Acquisition

G3- Individuals and Societies

G4- Sciences

G5- Mathematics

The strength of the relationship between two variables can be measured by a statistic called the correlation coefficient. A value of zero indicates no relationship between the two measures whereas a value of one indicates a perfect positive relationship. The table below shows the correlation coefficients between *CAT4* standard age scores and students' subsequent IB Grade outcomes at HL and SL levels. The correlation with the overall IB points is 0.35 with Mean *CAT4* SAS and 0.39 with Verbal SAS and these correlations are moderate. Correlations were highest with the Verbal section even for Mathematics. The correlations are low for Language Acquisition and moderate for other areas. The lower correlations compared with, for example, GCSE are partly due to the nature of the cohort as higher performing students tend to take the IB. For example, the mean *CAT4* SAS for the students taking Mathematics at SL level is 106 and the correlation is 0.32. The highest achieving candidates were generally most likely to take Mathematics at HL level as the mean *CAT4* SAS for these students is much higher at 115 and the narrower range of *CAT4* scores at the HL leads to the lower correlation of 0.21.

		Correlation	
Subject area	Level	Mean SAS	Verbal SAS
Individuals and Societies	HL	0.34	0.36
	SL	0.30	0.39
Language Acquisition	HL	0.17	0.22
	SL	0.24	0.21
Mathematics	HL	0.21	0.29
	SL	0.32	0.36
Sciences	HL	0.32	0.34
	SL	0.40	0.42
Studies in Language and Literature	HL	0.35	0.39
	SL	0.36	0.43
Overall IB points		0.35	0.39

The example below is taken from the Individual Report for Teachers for one student. It illiustrates the *CAT4* scores; the probabilities of obtaining each grade; the most likely grades and if challenged grades. The indicators are not precise, and this is reflected in the probabilities of obtaining each grade. These indicators will be updated as we collect more data.

CAT4 Individual student report for teachers   CAT4

Name: James Barros			
School: Test School			
Group: Year 10			
Date of test: 10/11/2011 – 10/11/2011	Level: F	Age: 14:02	Sex: Male

IB Diploma Programme indicators

Results from *CAT4* can give an indication of the grade a student will reach at the end of the Diploma Programme. A second grade is suggested – this is the grade a student could reach with additional effort and challenge. This information is helpful when you discuss with your students the targets they should be working towards.

Mean SAS: 105	Verbal SAS: 92	Quantitative SAS: 108	Non-verbal SAS: 106	Spatial SAS: 112

	Most likely	'If challenged'
Total IB Diploma points	29	32

	Probability of obtaining each grade					Most likely grade achieved	'If challenged' grade achieved	Probability of student obtaining grade 4 or higher / Probability of student obtaining grade 6 or 7
	1-3	4	5	6	7			10% 20% 30% 40% 50% 60% 70% 80% 90%
G2 Language acquisition HL	2%	10%	22%	43%	22%	6	7	
G1 Studies in language and literature HL	1%	22%	49%	23%	4%	5	6	
G1 Studies in language and literature SL	2%	31%	48%	17%	2%	5	6	
G2 Language acquisition SL	4%	21%	31%	34%	11%	5	6	
G3 Individuals and societies HL	5%	24%	43%	24%	5%	5	6	
G3 Individuals and societies SL	4%	22%	43%	25%	5%	5	6	
G4 Sciences HL	32%	30%	22%	11%	4%	4	5	
G4 Sciences SL	35%	30%	21%	10%	3%	4	5	
G5 Mathematics HL	19%	32%	29%	15%	5%	4	5	
G5 Mathematics SL	26%	35%	25%	11%	3%	4	5	

International Baccalaureate (IB) Pointers

IB pointers reports are available for the Middle Years and for Diploma level. An example of the individual student Middle Years report for parents is illustrated below. The pointers are based on the mapping exercise between GCSE grades in England and IB Middle Year grades (e.g. GCSE grade A*= 7 IB points, GCSE grade A = 6 points etc). The *CAT4* to GCSE grades have been estimated from a large dataset of students that have done both *CAT4* and GCSE (see section TBC). The IB pointers report at Diploma level is similarly based on a mapping exercise between A level grades in England and IB Diploma grades.

CAT4 Individual report for parents   CAT4

Name: James Barros			
School: Test School			
Group: Year 10			
Date of test: 10/11/2011	**Level:** F	**Age:** 14:02	**Sex:** Male

IB Middle Years Programme pointers

Subject	■ Most likely grade achieved	■ 'If challenged' grade achieved	IB Middle Years grade 1	2	3	4	5	6	7
Arts	5/4	5				■	■		
Mathematics	5/4	5				■	■		
Physical Education	5/4	5				■	■		
English	4	5				■	■		
French	4	5				■	■		
German	4	5				■	■		
Humanities	4	5				■	■		
Sciences	4	5				■	■		
Spanish	4	5				■	■		
Technology	4	5				■	■		

CAT4 trialling

Pre-trials

Small-scale trials were conducted in autumn 2009 to check some of the new questions being developed for the *CAT4* Spatial Ability Battery. Three versions of the new spatial test were created and were trialled with approximately 850 students in Years 4, 6, 8 and 9. Results from this study were used to develop further spatial questions for the main trials.

Main trials

The main trials of all the questions in all four batteries of *CAT4* were carried out in autumn 2010.

The numbers of students taking part in the trials were as follows:

Trial sample	
Year	Number of students
4	2,028
6	1,870
8	2,179
10	2,114
Total	8,191

For the trials, 24 test booklets were created, that is six test booklets for each year group. All students took Verbal Classification and Figure Recognition plus two of the remaining six test types, so that all items were taken by at least 300 students. Some of the questions were duplicated in booklets across year groups.

The data from the trials were analysed to provide information on the difficulty level of each question, its ability to discriminate between high and low scorers, and the extent to which it proved equally difficult for both sexes, once each sex's general level of performance was taken into account. This information was then used to select and order the sequences of questions for the final standardisation version of *CAT4*.

CAT4 UK standardisation: levels Pre-A to G

The standardisation of *CAT4* took place between September and December 2011 in England, Wales, Scotland and Northern Ireland. A national database of schools was created and schools were grouped into 10 categories – by country (Wales, Scotland and Northern Ireland) and, for England, further grouped into independent or grammar, plus five categories of school intake based on the proportion of students taking free school meals.

Schools were selected by stratified random sampling procedures within these groupings. As this was a national sample, many schools taking part in the standardisation had never used *CAT4* before. For the standardisation, schools were asked to do one pre-selected *CAT4* test level and were given an option to do other levels. Schools were free to choose between the paper and digital version of the test. Primary schools were asked to test all students in the year group but secondary schools had the option either to test two randomly selected teaching groups if they tested by paper, or to test the whole year group if they chose the digital option.

The numbers of students taking part in the standardisation were as follows:

	Standardisation sample		
Country	Primary	Secondary	Total
England	4,663	13,085	17,748
Wales	269	2,169	2,438
Scotland	259	2,439	2,698
Northern Ireland	179	1,645	1,824
Total	5,370	19,338	24,708

These numbers were compared with the national population:

	Standardisation sample			National population
Country	Primary	Secondary	Total	
England	87%	68%	72%	83%
Wales	5%	11%	10%	5%
Scotland	5%	13%	11%	8%
Northern Ireland	3%	9%	7%	3%
Total	100%	100%	100%	100%

Note: Totals may not add up to 100% due to rounding

The primary school sample is slightly over-represented by students from England and under-represented by students from Scotland. The secondary school sample is over-represented by students from Wales, Scotland and Northern Ireland and under-represented by students from England. The standardisation results were therefore weighted to account for sample bias.

The numbers of students doing the paper and digital editions are given below:

Delivery mode	Number of students in standardisation sample, by delivery method		
	Primary	Secondary	Total
Digital	1,123 (21%)	13,412 (69%)	14,535 (59%)
Paper	4,247 (79%)	5,926 (31%)	10,173 (41%)
Total	5,370	19,338	24,708

CAT4 standardisation: level X

Pre-trials

CAT4 Level X was developed after the main *CAT4* Levels A-G were published. Small-scale trials were conducted in Autumn 2009 to check some of the new Spatial questions being developed for CAT4. Three versions of the Spatial tests were created and were trialled with around 850 students in Years 4, 6, 8 and 9. Results from this study informed the development of further Spatial questions for trialling.

Main trials

The main trials of the *CAT4* Level X questions were carried out in Autumn 2013. Approximately 1200 students in Years 2 and 3 took part in the trials.

Four test booklets were created - two test booklets for each year group. Around 300 pupils took each booklet, with the parallel booklets of each year group alternated within a class. All the questions used in *CAT4* Level X were used in the trialling with some of the questions duplicated in booklets across the two different year groups.

The data from the trials were analysed to provide information on the difficulty level of each question, its ability to discriminate between high and low scorers and the extent to which it proved equally difficult for both sexes, once overall score was taken into account. This information was then used to select and order the sequences of questions for the final standardisation version.

Standardisation

The standardisation of *CAT4* Level X took place between May and June 2014 in England, Wales, Scotland and Northern Ireland. A national database of schools was created and schools were grouped into nine categories by country and within England. This was further grouped into 'Independent' plus five categories of maintained sector schools based on the proportion of students taking free school meals.

Schools were selected by stratified random sampling procedures within these groupings. As this was a national sample, many schools taking part in the standardisation had never used *CAT4* before. Around 1900 students completed Form X. The standardisation results were weighted to account for sample response bias.

The mean *CAT4* Level X standard age scores (SAS) for males and females are in the table below.

Gender		Non-verbal SAS	Verbal SAS	Quantita-tive SAS	Spatial SAS	Mean CAT4 score
Females	Mean	102.4	102.4	100.1	100.5	101.5
	N	944	941	941	941	945
	Std. Deviation	14.9	15.0	14.0	14.8	11.4
Males	Mean	98.6	98.6	100.3	99.3	99.2
	N	981	966	967	962	984
	Std. Deviation	14.6	14.8	16.9	15.0	12.0
Total including unknown	Mean	100.5	100.5	100.2	100.0	100.4
	N	1931	1913	1914	1909	1934
		14.8	15.0	15.5	14.9	11.8

Overall, female mean *CAT4* scores are around 2 SAS points higher than for males for Level X. In particular, the mean Verbal and Nonverbal scores are around 4 SAS points for females.

Note that the mean *CAT4* score is not a Standard Age Score but an average of the nonverbal, verbal, quantitative and spatial SAS. The standard deviation for the mean *CAT4* score is around 12, lower than the 15 that is expected for an SAS. This does not indicate the sample was unrepresentative in its spread of ability: rather, that the scores for the four components are correlated, so the spread narrows as scores are averaged.

CAT4 and teacher assessment levels

There is a significant and positive correlation between student's *CAT4* scores and their school performance, as measured by national tests or public examinations. The link may be assumed to exist because a lot of school activity is concerned with the application of reasoning abilities in the initial learning of curriculum content, and then building on and recombining existing knowledge as learning progresses.

During the standardisation, teachers in England and Wales were asked to provide information on students' current teacher assessment (TA) levels in English, Maths and Science for Level X.

The strength of a relationship between two measures can be expressed with a statistic termed a correlation coefficient. This coefficient goes from 0, indicating no relationship to 1 indicating a perfect relationship.

The table below illustrates the correlations between the *CAT4* standard age scores (SAS) and the TA levels. The mean *CAT4* score is the average of the verbal, quantitative, spatial and nonverbal reasoning SAS scores. The correlation coefficients are all highly significant. The figures in bold are the highest correlations for each test outcome. The

mean of the scores on all three batteries gives the highest correlations for Maths and Science. For English, the verbal battery gives a slightly higher correlation than the mean CAT4 score. Teachers reported sub-levels for English and Maths but reported whole levels for Science. The correlations with Science are slightly lower because the Science TA levels are reported as whole levels and hence do not discriminate as well.

Level X	Correlation		
	English level	Maths level	Science level
Non-verbal SAS	0.41	0.39	0.37
Verbal SAS	0.65	0.57	0.52
Quantitative SAS	0.47	0.48	0.40
Spatial SAS	0.42	0.43	0.43
Mean CAT4 score	0.63	0.61	0.55

Evaluating differences between *CAT4* scores

Evaluating a difference between two scores, whether scores on two different tests or scores on the same test on two occasions, has to be a three-stage process.

Statistical significance of differences

First, it needs to be decided if the difference is large enough to be considered as 'real' rather than just a result of having imprecisely measured the two scores. This depends upon the test reliability of each of the two scores and, hence, the 'noise' around each one.

The measurement error when calculating a difference between two scores is evaluated using a coefficient called the standard error of measurement difference (SEM_{diff}).

The SEM_{diff} for *CAT4* scores is approximately seven standard score points. Consequently, if two scores are more than seven SAS points apart, it is 68% likely that they are real, and if they are 11 points apart, the likelihood is 90% that the difference is a real one.

Rarity of differences

Second, if the difference is 'real' or statistically significant, then the **unusualness** or **rarity** of the difference has to be evaluated. A significant difference can sometimes be very common. For example, if you use a millimetre ruler to measure a boy's height when he is seven and then again when he is eight, the difference between these two heights can be measured very accurately to within two millimetres. Therefore 'real' or statistically significant differences will be very common in a sample of boys because the difference between the heights is likely to be substantially greater than two millimetres in almost all cases.

The spread of difference in scores can be determined either directly from the data or by a formula that takes into account the spread of scores on each test and the correlation between the two sets of scores. If the sample size is large enough, the two methods will produce very

similar results; this was the case for the standardisation of *CAT4*. The formula used is:

$$SEM_{diff} = \surd (SD_1^2 + SD_2^2 - 2r_{12} SD_1 SD_2)$$

where SD_1 and SD_2 are the standard deviations of the scores on each test and r_{12} is the correlation between the two tests.

When looking at differences between a child's scores on the same battery on two occasions (e.g. Verbal in Year 7 and Verbal in Year 8) the table below can be used[1]. For example, a score increase of 11 SAS points or more will occur with between 10% and 15% of children, but a decrease of 17 or more points will occur with only the most extreme 5%.

Difference in SAS scores from first to second occasion	Percentage of students obtaining this extent and direction of difference
Increases by >16	5%
Increases by >12	10%
Increases by >9	15%
Decreases by >9	15%
Decreases by >12	10%
Decreases by >16	5%

When looking at score differences between different batteries (e.g. Quantitative and Nonverbal), this table should be used instead[2]. The *SAS* score differences are larger in this situation because the two measures are of different underlying mental processes and so tend to be less highly correlated than two scores on the same test.

Difference in SAS scores from Battery 1 to Battery 2	Percentage of students obtaining this extent and direction of difference
Higher by >19	5%
Higher by >15	10%
Higher by >12	15%
Lower by >12	15%
Lower by >15	10%
Lower by >19	5%

[1] The figures in the table have assumed a mean correlation of 0.8 between the two occasions.

[2] The figures in the table have assumed a mean correlation of 0.7 between pairs of batteries.

Practical significance of differences

Finally, it needs to be remembered that a difference between two batteries which occurs commonly in the general population is not necessarily insignificant. It can indicate a real, albeit common, difference between the development of the cognitive abilities underlying the two battery scores, with implications for the ways in which the student concerned is likely to progress academically. Such differences need to be interpreted in the light of all that is known of a student's background and educational record. For example, students who have a background of poor socio-economic and educational opportunities who gain higher scores for Nonverbal Reasoning than for Verbal Reasoning may not have any real difference between their abilities to reason with words and with shapes. Instead, they may not have had the chance to acquire the basic reading and word knowledge needed to perform well on the verbal tasks. On the other hand, if they have good socio-economic and educational backgrounds, then the score difference may suggest that there is a genuine difference in abilities to think with words and with shapes.

Gender differences

The table below shows the mean SAS scores and standard deviation for each of the *CAT4* batteries and for primary and secondary schools. The results are based on 2,578 females and 2,792 males from primary schools; 9,471 females and 9,867 males from secondary schools.

School type	Gender		Verbal Reasoning SAS	Quantitative Reasoning SAS	Nonverbal Reasoning SAS	Spatial Reasoning SAS	Mean *CAT4* SAS
Primary	Female	Mean	100.8	99.3	100.1	99.4	99.9
		Std. Deviation	14.4	13.9	14.6	14.5	12.3
	Male	Mean	99.3	100.9	99.9	100.8	100.2
		Std. deviation	15.4	15.9	15.3	15.3	13.4
	Total	Mean	100.0	100.1	100.0	100.1	100.1
		Std. deviation	14.9	15.0	15.0	14.9	12.9
Secondary	Female	Mean	100.5	99.1	100.5	100.4	100.1
		Std. deviation	14.4	13.4	14.2	14.2	12.1
	Male	Mean	99.5	101.3	99.7	99.8	100.1
		Std. deviation	15.5	16.1	15.6	15.4	13.6
	Total	Mean	100.0	100.1	100.1	100.1	100.1
		Std. deviation	15.0	14.8	14.9	14.8	12.8

Verbal Reasoning scores in primary schools are on average around 1.5 SAS points higher for females than for males. In contrast, Spatial and Quantitative Reasoning scores are around 1.5 SAS points higher for

males than for females. There is not much of a gender difference for Nonverbal reasoning.

In secondary schools the Quantitative Reasoning scores are on average around two SAS points lower for females than for males. Average gender score differences for the other *CAT4* batteries are smaller – all within one SAS point.

The spread of scores as measured by the standard deviation is in general greater for males than for females. Therefore you are more likely to get proportionately more males than females having the extreme low or high SAS scores.

Verbal-Spatial profile

The table below shows the proportion of males and females within the verbal-spatial profile for primary and secondary schools.

Verbal-Spatial Profile	Primary			Secondary		
	Female	Male	Total	Female	Male	Total
Extreme spatial bias	1%	2%	1%	1%	2%	2%
Moderate spatial bias	3%	6%	5%	3%	6%	5%
Mild spatial bias	9%	11%	10%	9%	14%	11%
No bias	68%	67%	68%	66%	63%	65%
Mild verbal bias	13%	9%	11%	13%	10%	11%
Moderate verbal bias	5%	3%	4%	5%	4%	5%
Extreme verbal bias	1%	1%	1%	2%	1%	2%
	100%	100%	100%	100%	100%	100%

Note: Totals may not add up to 100% due to rounding

A total of 19% of females in primary schools have a verbal bias (mild, moderate and extreme categories) compared to 13% of males. In contrast, 19% of males in primary schools have a spatial bias compared with 13% of females.

A total of 20% of females in secondary schools have a verbal bias compared to 15% of males. In contrast, 22% of males in secondary schools have a spatial bias compared with 13% of females.

The gender difference among those with an extreme bias to spatial thinking are more striking. Overall, 2.3% of males show this profile, compared with only 0.8% of females. The bias is less differentiated by gender for those with an extreme bias to verbal thinking, with overall

1.8% of females and 1.3% of males being this category.

CAT4 standardisation for CBSE

A dedicated standardisation

Schools delivering the CBSE curriculum were interested in personalising learning and differentiating teaching by gaining an understanding of their students' learning profile and potential. *CAT4* is the ideal way of measuring these aspects of student development and offers both a means of monitoring cognitive development over time and indicators of future attainment.

The UK data for *CAT4* is a very good international benchmark and is used by schools across the world recognising the robustness and appropriateness of the data. However, when there is an opportunity to refine standardisation outcomes by collecting data from students and schools that can be considered homogenous and representative then this will be undertaken.

Working with a range of schools during 2014-15, *CAT4* was administered to a total of 12,864 students following the CBSE curriculum. Most of the testing was conducted in May and June 2015. The numbers of students and schools taking part by CAT level are given below.

CAT4 Level	School Grade	Number of students	Number of schools
Level B	Grade 4	4285	13
Level E	Grade 7	3411	11
Level F	Grade 9	3365	12
Level G	Grades 10-12	1803	5
	Total	12864	

Since the project started GL Education has begun working with other schools offering the CBSE curriculum across a number of countries and data collection is on-going allowing us to further refine test outcomes, principally the *CAT4* Indicators which have been updated as of 2019.

From the analysis of the standardisation results Standard Age Scores (SAS) were calculated. The SAS is based on the student's raw score which has been adjusted for age and placed on a scale that makes a comparison with a representative sample of CBSE students of the same age. The average score is set to 100. The SAS is key to benchmarking and tracking progress and is the fairest way to compare the performance of different students within a year group or across year groups.

The STEM agenda

The need to understand students' potential to achieve in the STEM subjects is global. The need in the UK to identify what could be lost potential in these subjects was one of the drivers for including a separate spatial ability section in *CAT4*. The opportunity to look at this

skills-set (alongside the other, more established, abilities) has appeal across many domains including in countries and in schools where the CBSE curriculum is offered.

The current project has achieved a standardisation based on students undertaking the CBSE curriculum. However, this is ongoing and GL Education is working with thousands of students in India and the Middle East to collect more data and refine its offering, particularly around Indicators for CBSE qualifications which have been updated as of 2019. GL Education is looking to collaborate with as many schools as possible and welcomes approaches from schools to take part in its programmes of testing.

The number of students taking part in the standardisation is given above.

Gender differences

The table below shows the average SAS scores for all the students who took part in the CBSE standardisation by gender.

Gender		Verbal SAS	Quantitative SAS	Nonverbal SAS	Spatial SAS	Mean *CAT4* SAS
Female	Mean	101.3	99.4	101.2	100.8	100.7
	Number of students	6788	6643	6826	6642	6641
Male	Mean	98.8	100.7	98.5	99.5	99.4
	Number of students	5995	5891	6037	5888	5885
Total	Mean	100.1	100.0	100.0	100.2	100.1
	Number of students	12783	12534	12863	12530	12526

Females are on average around 2.5 SAS points higher than males for Verbal and Nonverbal tests. For Spatial, females are on average around 1 SAS point higher than males, whereas for Quantitative females are on average around 1 SAS point lower than males.

Comparison with UK standardisation scores

CAT4 was originally nationally standardised in the UK in 2011. It is possible to compare the differences in SAS scores between an average CBSE and an average UK student. The table below shows these differences.

CAT4 Level	Average SAS differences* between UK and CBSE sample				
	Verbal SAS	Quantitative SAS	Nonverbal SAS	Spatial SAS	Mean difference
Level B	-6	-4	-3	-5	-4
Level E	-7	0	-3	-5	-4
Level F	-2	2	0	-2	-1
Level G	-2	2	0	1	0

Overall CBSE students score 4 points lower at levels B and E. At these *CAT4* levels Verbal shows the largest discrepancy, but this is expected as students are less familiar with the English language at these younger ages. The Verbal differences reduce significantly with age and by Grade 9 the difference is only 2 points. For practical purposes there are no significant differences between the UK and CBSE norms at levels F and G across the four *CAT4* test sections.

Data for the CBSE standardisation was collected for *CAT4* levels B, E, F and G. Given the consistency in score differences between UK and CBSE sample for levels B and E, we have used this information to estimate the CBSE standardisations for *CAT4* levels A, C and D. The standardisations will be updated as more data is collected.

CBSE Grade 10 indicators

CBSE Grade 10 exam results were collected from students who had done the *CAT4* test. Results were collected for

• English Communicative

• Hindi

• Mathematics

• Science

• Social Science

These indicators will be updated and more subjects added as we collect more data.

The strength of the relationship between two variables can be measured by a statistic called the correlation coefficient. A value of zero indicates no relationship between the two measures whereas a value of one indicates a perfect positive relationship. The table below shows the correlation coefficients between *CAT4* standard age scores and students' subsequent CBSE Grade 10 outcomes. These show that the overall Mean *CAT4* SAS has a moderate to strong association with the CBSE subject grades.

SAS	English Communicative	Hindi	Mathematics	Science	Social science
Mean CAT4	0.60	0.44	0.56	0.58	0.54

The example shows the grades most likely to be achieved by one student. The most likely grade for English Communicative is A1 (79 per cent) with an 18 per cent chance of obtaining grade A2, 2 per cent chance of obtaining grade B1. The indicators are not precise: they indicate the outcomes expected for students with a particular *CAT4* score making average progress in a typical school.

They come with a margin of error, which reflects the differences in progress that may be made by different students. This is reflected in the probabilities of obtaining each grade.

CAT4 Individual student report for teachers   CAT4

Name: Natasha Aransola

School: CBSE School

Group: EM

| **Date of test:** 21/10/2015 | **Level:** D | **Age:** 15:04 | **Sex:** Female |

CBSE Class X indicators

Results from *CAT4* can give an indication of the level a student will reach at the end of their secondary education. A second level is suggested – this is the level a student could reach with additional effort and challenge. This information is helpful when you discuss with your students the targets they should be working towards.

| **Mean SAS:** 117 | **Verbal SAS:** 107 | **Quantitative SAS:** 115 | **Non-verbal SAS:** 124 | **Spatial SAS:** 122 |

	Probability of obtaining each grade					Most likely grade achieved	'If challenged' grade achieved	Probability of student obtaining grade A2 or higher / Probability of student obtaining grade A1
	C1 or lower	B2	B1	A2	A1			10% 20% 30% 40% 50% 60% 70% 80% 90%
English Communicative	0%	0%	2%	18%	79%	A1	A1	
Hindi	1%	2%	6%	21%	71%	A1	A1	
Science	1%	2%	6%	23%	68%	A1	A1	
Social Science	1%	2%	6%	22%	70%	A1	A1	
Mathematics	2%	4%	9%	27%	59%	A2/A1	A1	

CBSE Grade 12 indicators

CBSE Grade 12 exam results were similarly collected from students who had done the *CAT4* test. Results were collected for Accountancy, Biology, Chemistry, Economics, English, Mathematics and Physics. These indicators will be updated and more subjects added as more data is collected.

The strength of the relationship between two variables can be measured by a statistic called the correlation coefficient. A value of zero indicates no relationship between the two measures whereas a value of one indicates a perfect positive relationship. The table below shows the correlation coefficients between *CAT4* standard age scores and students' subsequent CBSE Grade 12 outcomes. These show that the overall Mean *CAT4* SAS has a moderate to strong association with the CBSE subject grades.

SAS	Accountancy	Biology	Chemistry	Economics	English*	Mathematics	Physics
Mean *CAT4*	0.22	0.33	0.31	0.32	0.37	0.31	0.27

*For English the correlation shown is with Verbal SAS as it is slightly higher than for Mean *CAT4*

CAT4 IRISH EDITION

CAT4 Irish standardisation

Irish age-based norms for the *CAT4* were derived from the administration of four levels of the tests (D to G) to students in random samples of primary and second-level schools nationwide in 2012. The Irish version of the tests has the same content as the UK edition and is aimed at the following students:

Test level	Suitable for	Age range
Level D	5th and 6th classes	10:06-12:11
Level E	End 6th/First Year	11:06-13:11
Level F	Second/Third Year	12:06-15:11
Level G	Fourth/Fifth Year	14:06-17:00+

The numbers of students used in the Irish standardisations were as follows:

Test level	Number of students
Level D	1,733
Level E	1,818
Level F	1,678
Level G	1,387
Total	6,617

Test reliability

The reliability of a test is a measure of the consistency of a student's test scores over repeated testing, assuming conditions remain the same – that is, there was no fatigue, learning effect or lack of motivation. Tests with poor reliability might result in very different scores for a student across two test administrations.

The test reliabilities of the Irish version are high and are similar to the UK edition.

Test level	*CAT4* reliability				
	Verbal Reasoning Battery	Quantitative Reasoning Battery	Nonverbal Reasoning Battery	Spatial Ability Battery	Overall *CAT4*
Level D	0.89	0.90	0.88	0.87	0.96
Level E	0.89	0.88	0.86	0.87	0.95
Level F	0.90	0.87	0.84	0.88	0.95
Level G	0.91	0.86	0.83	0.88	0.95
Average D-G	**0.90**	**0.88**	**0.85**	**0.87**	**0.95**

For interpreting the score of an individual student, the standard error of measurement (*SEM*) is a more useful statistic than a reliability coefficient. It indicates how large, on average, the fluctuations in standard scores may be. The *SEM* for the Verbal Reasoning Battery is 4.8, which indicates that there is a 68% chance that the student's true verbal SAS will be in the range +/- 4.8. For example, for an average-performing student with a verbal SAS of 100, there is a 68% chance that his or her true verbal score is in a range from 95 to 105.

CAT4 level	CAT4 standard error of measurement (*SEM*)				
	Verbal Reasoning Battery	Quantitative Reasoning Battery	Nonverbal Reasoning Battery	Spatial Ability Battery	Overall CAT4
Average D-G	4.8	5.3	5.8	5.3	3.0

However, most tests show the 90% chance or confidence bands. For values around the average, the 90% confidence band is as follows:

CAT4 level	CAT4 90% confidence band				
	Verbal Reasoning Battery	Quantitative Reasoning Battery	Nonverbal Reasoning Battery	Spatial Ability Battery	Overall CAT4
Average D-G	+/- 8	+/- 9	+/- 9	+/- 9	+/- 5

For example, for an average-performing student with a verbal SAS of 100, there is a 90% chance that the true verbal score is in a range from 92 to 108.

Gender differences

The table below shows the average SAS scores for all the students who took part in the Irish standardisation, by gender.

Gender		Verbal Reasoning SAS	Quantitative Reasoning SAS	Nonverbal Reasoning SAS	Spatial Reasoning SAS	Mean CAT4 SAS
Female	Mean	100.4	98.9	100.6	99.5	99.8
	Number of students	3,750	3,745	3,750	3,715	3,766
Male	Mean	99.6	101.9	99.6	101.1	100.4
	Number of students	2,714	2,700	2,719	2,676	2,737
Total	Mean	100.0	100.1	100.2	100.2	100.1
	Number of students	6,574	6,556	6,578	6,499	6,617

Males were on average around three SAS points higher and around 1.5 SAS points higher for Spatial. Females were around one SAS point higher than for males the Verbal and Nonverbal Batteries.

Irish Leaving Certificate indicators

Results were collected from 870 students who completed *CAT4* and the Leaving Certificate. Subject grades were obtained as either Ordinary (O) or Higher (H) level. The equivalence between the Ordinary and Higher grades as set out in https://www.cao.ie/index.php?page=scoring&s=lcepointsgrid was used to combine results from the two levels to a common scale. For example, Higher 6 grade is equivalent to Ordinary 2 grade and both of these have 46 points.

The strength of the relationship between two variables can be measured by a statistic called the correlation coefficient. A value of zero indicates no relationship between the two measures, whereas a value of one indicates a perfect positive relationship. The correlations between *CAT4* scores and Leaving Certificate subjects grades are shown below. These show that the overall mean *CAT4* SAS has a moderate to strong association with the subject grades.

	Mean *CAT4* score	Verbal SAS	Quantitative SAS	Nonverbal SAS	Spatial SAS
Art	0.55	0.52	0.42	0.48	0.47
Biology	0.60	0.63	0.43	0.48	0.44
Business	0.50	0.60	0.42	0.30	0.30
Chemistry	0.52	0.50	0.48	0.40	0.29
Construction Studies	0.52	0.44	0.31	0.40	0.45
English	0.58	0.67	0.42	0.43	0.35
French	0.54	0.59	0.39	0.41	0.35
Geography	0.60	0.63	0.42	0.43	0.45
History	0.46	0.52	0.35	0.30	0.33
Home Economics	0.43	0.54	0.42	0.32	0.20
Irish	0.40	0.46	0.29	0.31	0.23
Maths	0.65	0.57	0.53	0.53	0.47
Physics	0.53	0.50	0.46	0.44	0.38

The Leaving Certificate indicators for each subject are derived from the statistical relationship between *CAT4* scores and Leaving Certificate subject grades or points scores. Indicators are calculated from the mean *CAT4* Standard Age Score (SAS) for Maths, Physics, Chemistry, Art and Construction Studies and are based on verbal SAS for the other subjects.

Likelihood of Leaving Certificate grades

The example below shows the grades most likely to be achieved by one student. The most likely grade for Construction Studies is H3 but this student has an 18% chance of obtaining a grade higher than H3, a 26% chance of obtaining grade H4 and a 23% chance of obtaining a grade below H4. The indicators are not precise: they indicate the outcomes expected for students with a particular *CAT4* score making average progress in a typical school. They come with a margin of error which reflects the differences in progress that may be made by different students. This is reflected in the probabilities of obtaining each grade.

CAT4 and Leaving Certificate 'Best 6' score

A summary 'Best 6' indicator based on the total points score for Maths and the best of five other subjects was calculated for each student. The correlation between the 'Best 6' points score and the mean *CAT4* score was 0.61 and the relationship is displayed graphically below.

Relationship between Best 6 points score and mean *CAT4* score

Mean *CAT4* score
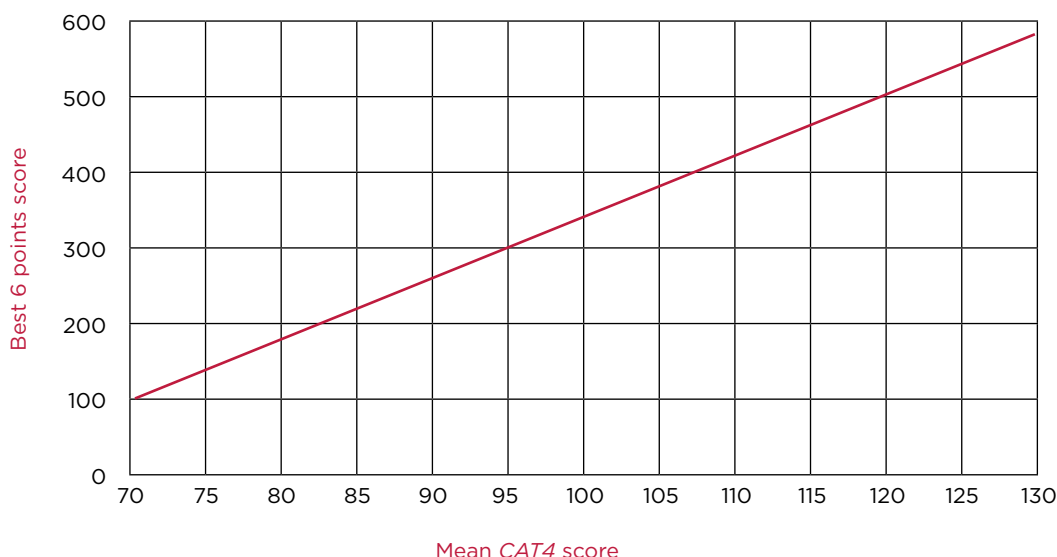
Leaving Certificate indicators for groups of students

The table below shows how the group/class indicators have been calculated for a fictitious class with five students and shows the most likely grade achieved and the probabilities associated with getting different Mathematics grades. The group indicator is an average of the individual student outcomes and probabilities.

Calculating group indicators for Mathematics for a fictitious class of five students

Student	Mean *CAT4* score	'Best 6' score	Leaving Certificate grades – probabilities									Most likely grade achieved
			<=O5	O4	O3	H6/O2	H5/O1	H4	H3	H2	H1	
1	70	99	95%	3%	1%	0%	0%	0%	0%	0%	0%	O5
2	85	217	72%	14%	8%	4%	1%	1%	0%	0%	0%	O5
3	100	337	25%	20%	21%	16%	8%	5%	3%	2%	0%	O3
4	115	457	4%	5%	10%	17%	16%	17%	17%	11%	3%	H5/O1
5	140	600	0%	0%	0%	1%	2%	3%	10%	36%	47%	H2
Group indicator (average)		342	39%	8%	8%	8%	6%	5%	6%	10%	10%	