

What CAT4 tells you

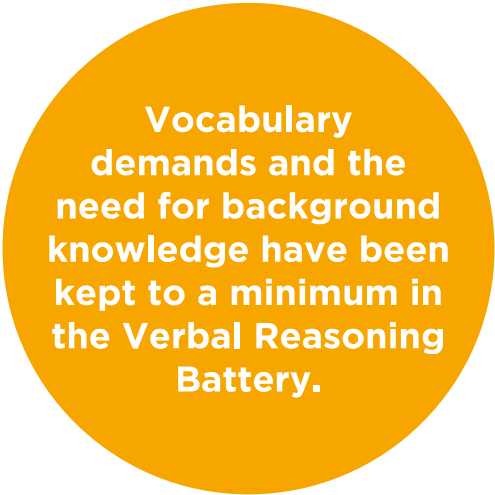
The four batteries of CAT4 assess a student's ability to reason with different kinds of material and so provide information that is highly valuable for both understanding students' strengths and diagnosing their learning needs.

What the four batteries assess

The Verbal Reasoning Battery assesses reasoning ability with words representing objects or concepts. The tests in this battery do not focus on the physical properties of the words themselves, such as the alphabetical position of their first letters. Likewise, the Quantitative Reasoning Battery assesses reasoning with numbers, with the numbers representing the relevant numerical concept, rather than being used for their physical properties such as whether they consist of two digits or one. The Nonverbal Reasoning and Spatial Ability Batteries are somewhat different in that the shapes themselves are the focus of the assessment rather than the shapes symbolising something else.

Thinking with words

The Verbal Reasoning Battery necessarily requires some reading ability. However, CAT4 limits the reading requirements to a modest level throughout. The vocabulary demands of the Verbal Analogies and Verbal Classification tests have been kept as low as possible. Also, the background knowledge needed to answer the verbal questions is that which all students will have encountered in school or everyday life, rather than including topics that may only be familiar to certain socio-economic or cultural groups.



Vocabulary demands and the need for background knowledge have been kept to a minimum in the Verbal Reasoning Battery.

Consequently, scores on the Verbal Reasoning Battery will usually reflect students' ability to use words as a medium of thought. The exceptions will be when students have poor reading skills or grew up apart from mainstream UK society.

It is also worth noting that all the instructions for the CAT4 batteries are presented orally to students, so any influence of reading skills is limited solely to the items in the Verbal Reasoning Battery.

Thinking with numbers

The Quantitative Reasoning Battery has been designed to be minimally reliant on mathematical knowledge. The Number Analogies test requires only basic arithmetical knowledge, and parallels the analogy tests in the Verbal and Nonverbal Reasoning Batteries. The Number Series test focuses as far as possible on the identification of relationships between the elements of the questions, though basic arithmetical knowledge is necessarily required too.

Mathematical knowledge is kept to a minimum in the Quantitative Reasoning Battery, although basic arithmetic is needed.

In this way, the Quantitative Reasoning Battery will give a genuine indication of most students' ability to think with numbers, with the exception of children with particularly low arithmetic skills.

Thinking with shapes

The Nonverbal Reasoning Battery assesses the ability to think and reason with Nonverbal material, that is to analyse figures made up of multiple elements, identify the relationships between these elements and identify further examples of these relationships. The Figure Matrices test parallels the analogies tests in the Verbal and Quantitative Reasoning Batteries. The Figure Classification test requires the identification of common elements between figures and parallels the Verbal Classification test.

The Nonverbal Reasoning Battery does not rely on high level verbal skills or English.

Consequently, the Nonverbal Reasoning Battery reveals how well students can think when working with shapes. As these questions do not necessarily rely on highly developed verbal skills or the use of English for their solution, they can provide insight into the reasoning abilities of students with poor verbal skills or who are not particularly fluent in English.

Caution may need to be exercised when interpreting low scores if the student concerned comes from a non-Western cultural background, as he or she may not have experienced these types of activities before.

Thinking about shape and space

The Spatial Ability Battery assesses the ability to think in spatial terms, that is to visualise shapes and objects and the effects of manipulations on these. The Figure Analysis test requires the student to imagine the effect of a series of physical manipulations on a square of paper. This test relies on both spatial and reasoning abilities, such as recognising that, if a hole is made through layers of a doubled-over sheet, there must be two holes when the sheet is unfolded. The Figure Recognition test requires the identification of a target shape within a complex design, so assessing the ability to identify a remembered shape from within more complex information.

Students with a high spatial ability may be well-suited to jobs involving visual mapping such as architecture, graphic design, photography and astronomy.

As spatial tests make no demands on verbal ability, they can be highly effective indicators of potential in students with poor verbal skills, as well as effectively identifying the weaker abilities of those who have verbal strengths. This then provides a more comprehensive picture of the students concerned.

As with the Nonverbal Reasoning Battery, caution needs to be exercised when interpreting low scores if students come from non-Western cultural backgrounds, owing to their potential lack of familiarity with this type of activity.

Scores from CAT4

For each CAT4 test students obtain a raw score which indicates the number of questions they answered correctly.

These raw scores are interpreted by comparing them to the performance of other students of the same chronological age group using what are referred to as 'normative scores'. Three types of normative score are provided for the interpretation of performance: Standard Age Scores (SAS); National Percentile Rank (NPR) by age; and stanines (ST) by age.

 **Standard Age Scores (SAS):** These are presented on a standardised score scale where the average for each age group is set to 100 and the standard deviation set to 15.² This means that a student who gains the same SAS on two different batteries has done equally well on both, compared to others of the same age. It also means

² This means that approximately 68% of students in the norm group for that age scored between 85 and 115, approximately 95% scored between 70 and 130, and over 99% scored between 60 and 140, the limits of the CAT4 SAS.

that students of different ages who have the same SAS have done equally well when judged in relation to others of their own age.

✿ **National Percentile Rank (NPR):** This indicates the proportion of students of the same age who have scored the same as or below the student in question. For example, a student who achieves a percentile rank of 84 has scored equal to or better than 84% of students in the same age band; only approximately 16% of students achieved a higher score on this test.

✿ **Stanines (ST):** This is a standardised score scale divided into nine bands. In a stanine scale the scores are grouped as shown in the table below. Stanines are particularly useful when reporting test results to students and parents as they are relatively easy to understand and interpret. They also avoid the erroneous impression of being 'IQ scores', sometimes attributed to SAS.

The stanine scale				
	Stanine	Percentage of cases	Corresponding percentiles	Corresponding SAS
Very high	9	4%	97 and above	127 and above
Above average	8	7%	90-96	119-126
	7	12%	78-89	112-118
Average	6	17%	59-77	104-111
	5	20%	41-58	97-103
	4	17%	23-40	89-96
Below average	3	12%	12-22	82-88
	2	7%	5-11	74-81
Very low	1	4%	4 and below	73 and below

CAT4 levels X and Y report SAS scores with limits different to CAT4 levels Pre-A to G.

While levels X and Y have limits of 69 – 131, levels Pre-A to G have limits of 59 – 141. This is because CAT4 levels X and Y are less reliable at measuring extremely good or extremely bad performances, because the tests are shorter at these levels. The tests are designed to be shorter because young children tend to have shorter concentration spans.

The SAS range that is common to levels Pre-A to G and X and Y function the same. Indeed, both mean and standard deviation is the same. Very few students are affected by the different limits to score reporting.

In order to understand the difference between CAT4 Levels X and Y standard age score reporting and the other CAT4 levels please see further guidance at: <https://www.gi-assessment.co.uk/content-pages/cat4-x/>

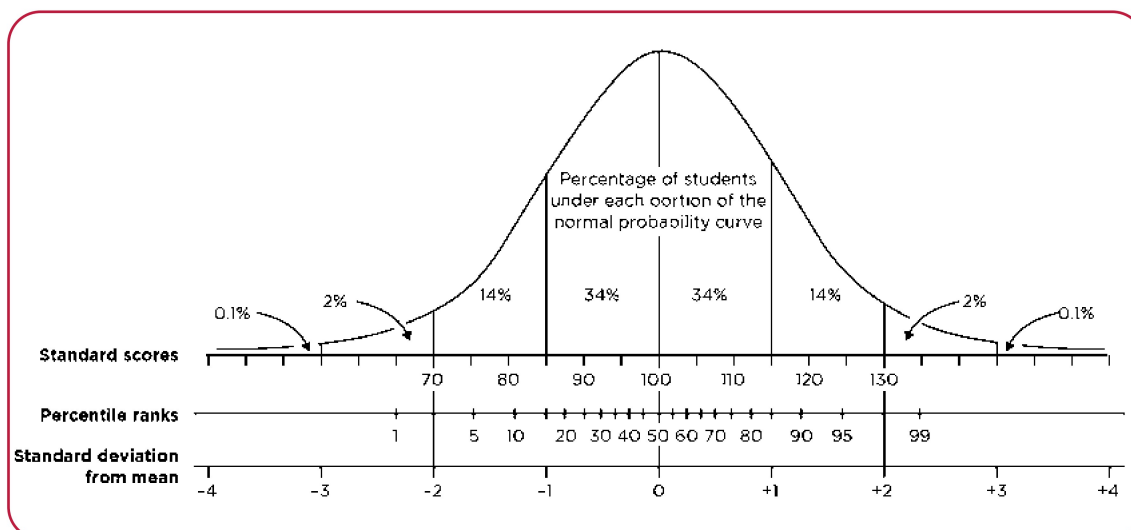
Relationship between CAT4 scores

The relationship between the three types of normative score is shown below, along with the normal distribution curve which illustrates the distribution of test performance in each age range.

Figure 1: Relationship between scores

Description	Very Low		Below Average		Average			Above Average		Very High			
Stanine (ST)	1	2	3	4	5	6	7	8	9				
Standard Age Score (SAS)	70	80	90	100	110	120	130						
National Percentile Rank (NPR)	1	5	10	20	30	40	50	60	70	80	90	95	99

Figure 2: Normal distribution curve



Low or unreliable scores

If a student's score on any one of the batteries is very low, it should be regarded with caution.

Before interpreting an individual student's score on any of the CAT4 reports, scan the report and find the number of questions attempted. This will show if a student has left a large number of questions unanswered on any of the batteries, or if his or her score is close to that expected from random guessing.

Examples of low and potentially unreliable scores are illustrated in the case studies found in this pack.

If all or nearly all of the questions have been attempted, then random guessing will result in raw scores at the 'chance level' shown in the table below. If fewer questions have been attempted then random guessing will, on average, result in a raw score of around one-fifth of the number of questions attempted.

This table shows chance levels of performance and these should be used to identify any students whose scores should be looked at more closely.

If the raw score is the same as or lower than the chance level given for the battery, then caution should be exercised in interpreting the score.

Chance levels of performance on CAT4 batteries		
	Maximum raw score	Chance raw score
Verbal Reasoning Battery	48	10
Quantitative Reasoning Battery	36	7
Nonverbal Reasoning Battery	48	10
Spatial Ability Battery	36	7

Any student who omits a large number of questions, or answers most of the questions but gets few of them right, is probably functioning at a low level in the cognitive area being tested. In either case, the student's score cannot be relied upon with confidence. Although these scores might actually represent the true level of the student's abilities at the time of testing, a better view of what the student can do might be obtained by retesting with CAT4 after a gap of at least six months – consider assessing the student with a series of tests that look at ability, processing and attainment, which might point to a specific learning difficulty; or seek outside support from an educational psychologist who can carry out a specialist assessment.