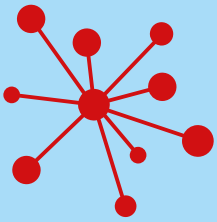
GL Assessment®

GL Education®

ADMINISTRATION MANUAL



Recall

Specialist Consultant
Dr Helen St. Clair-Thompson

COMPUTERISED ASSESSMENT OF WORKING MEMORY AND PROCESSING SPEED FOR 7 TO 16 YEAR OLDS

Third Edition, 2025

GL Assessment Limited is a GL Education Group Limited

Typeset by Mike Connor Design & Illustration

Main distributors:
GL Assessment
Vantage West, Great North Road
Brentford, TW8 9AG, UK

www.gl-assessment.co.uk

Contents

1 Introduction

1.1 What is Recall?

Recall is a suite of computerised tests designed to assess working memory skills in the age range 7 years 0 months to 16 years 11 months. The Recall suite comprises standardised tests of the following memory processes:

- Phonological loop (Word Recall test)
- Visuo-spatial sketchpad (Pattern Recall test)
- Central executive function (Counting Recall test)

In addition, Recall provides the following additional standardised measures derived from those core tests:

- Composite working memory skills
- Working memory processing speed

Test administration is carried out entirely online through the Testwise platform.

Each test begins with spoken instructions and practice items. The total suite takes 20–30 minutes. Full details of the tests in Recall, including guidelines on test administration, are given in Chapter 2. Results, based on nationally standardised norms, are available after the tests are completed. Results are given in standard age score (SAS) and centile score formats and age equivalents within the age range 5:0 – 16:11. Guidance on understanding results and interpreting reports are given in Chapter 3, with advice on how children with poor working memory can be helped given in Chapter 4. Finally, Chapter 4 discusses a number of illustrative case studies, providing pointers for effective intervention and classroom support.

1.1.1 What is working memory?

The distinction between short-term memory and long-term memory is widely understood. Short-term memory refers to remembering a small amount of information for 15-30 seconds, and long-term memory refers to the vast amount of information that we acquire over a lifetime. Working memory is an alternative term for short-term memory, capturing that we need to hold information in mind temporarily during the completion of cognitive tasks. Imagine, for example, being given a set of instructions for a task. You need to use working memory to remember the instructions whilst carrying them out. Now imagine multiplying two numbers together. The numbers need to be held in working memory whilst performing the calculation. Similarly, during reading comprehension text must be remembered whilst it is processed to uncover its meaning. Working memory is therefore involved in many everyday tasks.

Scientists have proposed different theories about the structure of working memory. However, there is general agreement that working memory is best understood as comprising storage

components along with a control or processing component. The storage components temporarily store information, whereas the control component is sometimes likened to attention, and is responsible for processing information and retrieving information from long-term memory. One such theory of working memory, which has been widely applied in educational settings, was proposed by Baddeley (Baddeley, 2000; Baddeley & Hitch, 1974). There are two domain-specific storage components; the phonological loop that is responsible for the maintenance of auditory information, and the visuo-spatial sketchpad that is specialised for dealing with visual and spatial information. These are governed by a control component called the central executive system. This controls the flow of information to and from the storage components and is responsible for processing or manipulating information. There is substantial evidence for this multiple-component model, which has come from experimental studies in cognitive psychology, the study of brain-damaged patients, and investigations of the brain areas that are active during working memory tasks. It is this theory of working memory which informed the development of Recall.

1.1.2 Why is working memory important?

Working memory plays an important role in learning in the school classroom. Many learning activities require children to hold in mind some information (e.g., a sentence to be written down) whilst engaging in a related activity (e.g., spelling the words). Classroom tasks require children to follow instructions to do one thing after another. Children also need to remember where they are up to when they are engaged in a task (e.g., which word in the sentence they need to spell next, or which instruction they now need to follow). Children with poor working memory will struggle in these tasks, often failing to complete learning activities because they have forgotten the information that was needed.

A substantial body of scientific research has examined the role of working memory in children's learning, revealing that working memory predicts performance in literacy, comprehension, mathematics, second-language learning, and science (e.g., Alloway, 2006; Cowan, 2014; Gathercole & Alloway, 2008; Sana & Fenesi, 2025). Between the ages of 7 and 14 years, children who perform poorly on measures of working memory also typically perform below expected standards in national curriculum assessments of English, mathematics, and science carried out in England (e.g., Gathercole et al., 2003; Gathercole et al., 2004; St Clair-Thompson & Gathercole, 2006).

Given how important working memory is for children's learning and success in education it is essential that we create a school environment that supports children with poor working memory. This includes identifying children who have poor working memory and adapting teaching practices. We need to be mindful of tasks that are very demanding of working memory, try to reduce these demands, and provide appropriate support structures. This may include considering adjustments to examination conditions. Recall can assist with this.

1.1.3 The diagnostic value of working memory assessment

Consistent with the idea that working memory is a key cognitive skill for children's learning, working memory difficulties are known to be associated with a range of learning difficulties and neurodevelopmental disorders. Children with reading difficulties typically perform poorly

on working memory tasks, particularly those that require the storage of verbal information (employing the phonological loop), and storage of verbal information whilst processing (also requiring the central executive). Similarly, children with dyslexia or specific language impairments tend to perform poorly on these measures. Children with mathematical difficulties tend to have poorer scores on tasks requiring the storage of visuo-spatial information (employing the visuo-spatial sketchpad), and storage of information whilst processing (also requiring the central executive). Working memory difficulties have also been identified in neurodevelopmental disorders, including for example Attention Deficit Hyperactivity Disorder, Autism Spectrum Disorder, Down syndrome, and Williams syndrome (Alloway & Gathercole, 2006b). Many children with neurodevelopmental disorders have difficulties in the school classroom. Performance on the tests in Recall can therefore be used to help understand the nature of a student's educational difficulties, and sometimes to give advance warning of likely difficulties as they get older, enabling early intervention measures to alleviate the educationally disadvantaging effects of memory limitations. Advice on this can be found in Chapter 4, with case studies illustrating these principles in action being provided in Chapter 5.

1.1.4 The need for Recall

There are several assessments of working memory which could be used in educational settings. Some well-known psychological test batteries, such as the Wechsler Intelligence Scale for Children (WISC) and the British Abilities Scales (BAS), include working memory measures – most typically forwards and backwards digit recall. There are also several comprehensive assessments of working memory, such as the Working Memory Test Battery for Children (Pickering & Gathercole, 2001) and Automated Working Memory Assessment, 2nd Edition (AWMA-2) (Alloway, 2012). However, most of the existing measures are designed for individual administration and require extensive teacher or assessor time. Furthermore, administration of the WISC and the BAS is restricted to appropriately qualified psychologists. These factors severely restrict the utility of other existing assessment products. The construction of Recall was motivated by the absence of a brief assessment of working memory which can be readily used in schools, is easy to administer, and is fully automated so that it does not require teacher or assessor input. Computer-based tests meet these requirements and also offer additional advantages, as outlined in the next section.

1.1.5 Use of Recall in examination access arrangement assessments

Recall results may also be used as evidence for examination access arrangements as a standardised assessment of:

- speed of reading; or
- speed of reading comprehension; or
- speed of writing; or
- cognitive processing measures which have a substantial and long term adverse effect on speed of working.

Assessors planning to use Recall for this purpose should be fully familiar with current JCQ regulations, which stipulate the qualifications of assessors and conditions for assessment. In particular, the regulations state that the assessment must be carried out by a suitably qualified person, who could be a psychologist or a specialist teacher, and the Head of Centre must satisfy themselves that this person is competent to carry out such assessments. This person then takes responsibility for selecting appropriate tests, interpreting the results, and making the recommendations for access arrangements. Careful administration is advised when using group assessment in order that individual student responses are observed and monitored. JCQ regulations give guidance on what qualifications and experience may be expected of named specialist teachers.

1.1.6 Advantages of computerised tests

The use of computerised cognitive assessments as an alternative to traditional paper-and-pen tests has gained momentum since the 1990's, with several assessments for working memory being available. In more recent years, remote assessments have also become increasingly popular, owing partly to fast and low-cost internet. One of the great advantages of a well-designed online test is that it does not require any special expertise or training on the part of the teacher or administrator. Digital tests also provide more precise measurement, especially when complex cognitive skills are being assessed. Tests are administered in an entirely consistent manner for all persons taking the test, which enhances reliability of measurement. Timings and presentation speeds can be controlled precisely. The subjective judgment of the teacher or administrator does not affect the test outcome as it may in conventional tests. Recall is largely self-administered and results are available shortly after the test; both of these factors help to reduce administrative load and avoid time delays. Provided headphones are used and certain basic precautions are taken, Recall can be administered in a room where other activities are taking place. For further information about test administration, see Section 2.2.

There is good evidence that most students prefer computer-based tests to conventional tests (whether paper-based group tests or administered 1:1 by a teacher or psychologist). This is particularly the case for less able students or those with below average literacy skills, who are more likely to feel intimidated by assessments and be embarrassed by their performance. Computer-based tests have generally been found to be less threatening and less stressful, which helps to ensure more reliable results (Singleton, 2001). There is also evidence that there is less gender bias in computer-based tests than in conventional tests, so there are good reasons to regard computer-based tests as fairer, as well as being more consistent and objective, than conventional tests (Horne, 2007).

When using paper tests, retesting can be problematic because such tests typically have fixed item order and content. When encountering the test for a second or subsequent time, students may remember items and answers, which may enable them to improve their performance over previous attempt(s). There may also be increased confidence from being confronted by familiar tasks rather than novel tasks (although students who prefer the excitement and challenge afforded by new and unfamiliar tasks may actually find this demotivating). These are usually referred to as practice effects, and in order to reduce practice effects it is generally recommended that there should be a suitable time interval between testing and retesting so that recollection is sufficiently diminished. Some test manuals advocate at least 12 months

between assessment, while others suggest a less stringent 3–6 months. A particular advantage of computerised tests (including Recall) is that test items can be generated randomly or drawn from a large item bank so that, although the student will be confronted by the same task on retesting, the items will be different from last time, thus reducing practice effects. This means that no minimum time interval needs to be placed on retesting, nor on the number of occasions that retests are given. This can be very useful when needing to evaluate the impact of an intervention over time.

1.2 Development of Recall

1.2.1 Test development

As noted in section 1.1.1., the choice of measures to include in Recall was guided by the multiple component model of working memory (Baddeley, 2000; Baddeley & Hitch, 1974). One core test (Word Recall) was included to assess the phonological loop. A large body of research has identified immediate serial recall as a paradigm that is suitable to assess phonological loop functioning. Another core test (Pattern Recall) was included to assess the visuo-spatial sketchpad. This provides a measure of visual short-term memory (Della Sala, Gray, Baddeley, & Wilson; 1997). The final core test (Counting Recall) was included to assess the central executive component of working memory. This is a variant of a complex span task in which participants must simultaneously store and process information (e.g. Case, Kurland & Goldberg, 1982). It is assumed that such tasks involve both the storage and central executive components of working memory. All these tasks have been used extensively in scientific research examining the role of working memory in learning and educational attainment, although there have been many variations in terms of how the tasks have been presented, administered, and scored.

A description of each test is provided in Section 2.1. Each of the three core tests is adaptive, with progress through each test and the point of discontinuation being determined by cumulative performance. More able students will progress through the test quicker and reach higher levels. Less able students will progress more slowly and generally avoid the unnecessary frustration of levels that are much too difficult for them. This reduces assessment time and helps to maintain the test-taker's motivation regardless of ability.

1.2.2 Types of results

Normative results in standard age score (SAS) and centile score form are incorporated into the Recall program for each of the three core measures. The norms are provided for the ages 7:0 to 16:11, together with confidence intervals and age equivalents. These different types of results are explained in Chapter 3.

An overall measure of general working memory functioning, called Working Memory Composite, was obtained by combining the scores of the three subtests with appropriate weighting to allow for differential item length.

A measure of processing speed was derived from the Counting Recall test by means of an algorithm that reflects the average time taken to count each item. This derived measure is referred to as Working Memory Processing Speed.

For each of the three core tests comparative results are provided for Memory Span (based on the maximum difficulty level reached in each subtest). Memory span is a measure of the number of items of information that the person can hold in memory at any given time.

The reason why standard scores for memory span have not been provided is explained in Section 1.3.3.

1.3 Standardisation

1.3.1 What is standardisation?

Technically, 'standardisation' is the process used in psychometric test development to create norms so that the performance of students of different ages can be represented by means of scores that are independent of age. However, the term 'standardised' is sometimes used in a non-technical sense to refer to the consistent administration of a test – i.e. that test instructions and methods of administration are the same for all who take the test. Because this non-technical usage can be misleading (e.g. users may assume that a test has standardised norms when in fact it hasn't) we only use the terms 'standardisation' or 'standardised' in strict accordance with technical psychometric usage.

The most common normative scores are standard scores (SAS) and centile scores. Standard scores have a mean (average) of 100 and a standard deviation (abbreviated to SD) of 15. Centile scores (sometimes known as percentile scores) place individuals on a 'ladder' of attainment from 1 to 100 compared with the population of that age; e.g. a centile score of 70 means that 70% of people would have lower raw scores and 30% would have higher raw scores. (For further information about standard scores and centile scores see Section 2. The standard deviation is the most common statistic for expressing variability in a set of scores and is calculated as the average amount by which the scores in the set deviate from the mean.

Thirty schools were recruited for the standardisation process. The schools were selected to include the age range of 7-16 years, and included both urban and rural schools to give a representative spread of school types and socio-economic profiles. Students were taken on an unselected basis from entire classes of students in the participating schools. No students were excluded from taking part on any basis.

1.3.2 Standardisation sample

The standardisation sample comprised 4843 students aged 7-16 years (2263 males and 2580 females) [see Table 1].

1.3.3 Standardisation results

All raw data from the three tests and also the two derived measures (Working Memory Composite and Processing Speed) approximated to normal distributions (symmetrical bell-shaped curves), with skewness (the degree of asymmetricality of the distribution) and kurtosis (the degree of flatness and peakedness of the distribution) below the critical threshold of 1.0. Descriptive statistics for each of the core tests are given in Table 2, and for the two derived measures in Table 3. For the three core tests and Working Memory Composite the developmental progression in raw score means from the youngest to the oldest age group is approximately linear with the exception of the 16:0-16:11 age group. From Table 1 it can be seen that the number of students in the 16:0-16:11 age group was significantly smaller than the other groups, and this is the most likely explanation for the divergent results pattern found in this group. For Processing Speed, the curve is approximately linear in the range 7:0-12:11, but plateaus thereafter, as might be expected with a speed measure.

Table 1. Number of students in the standardisation sample by age.

Age	Males	Females	Total
7:0 - 7:11	201	212	413
8.0 - 8.11	302	301	603
9:0 - 9:11	300	346	646
10:0 - 10:11	311	275	586
11:0 - 11:11	355	313	668
12:0 - 12:11	343	250	593
13:0 - 13:11	288	207	495
14:0 - 14:11	322	229	551
15:0 - 15:11	132	120	252
16:0 - 16:11	26	10	36
All (7:0 - 16:11)	2580	2263	4843

Table 2. Raw score means and standard deviations for the tests in Recall by age.

Age band	Word Recall		Pattern Recall		Counting Recall	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
7:0 - 7:11	8.74	4.98	21.42	8.89	7.29	6.24
8.0 - 8.11	10.78	5.28	23.11	8.48	10.32	6.25
9:0 - 9:11	11.59	4.96	28.75	8.59	13.04	7.13
10:0 - 10:11	13.52	5.03	32.68	8.74	14.88	6.88
11:0 - 11:11	14.47	5.15	36.80	8.79	15.61	7.46
12:0 - 12:11	15.58	5.01	37.10	9.10	15.43	7.70
13:0 - 13:11	16.10	5.30	39.58	9.93	16.34	8.07
14:0 - 14:11	17.24	5.59	40.37	11.75	18.72	8.38
15:0 - 15:11	18.38	4.82	42.17	10.37	22.44	8.31
16:0 - 16:11	16.51	4.18	40.56	13.15	20.82	7.68

The overall breakdown of data was considered appropriate for standardisation; however, norms in the age range 16:0-16:11 should be regarded as provisional for the time being because the number of students in this age range fell below psychometric conventions. The norms for this age were adjusted using extrapolated scores from the development curve for ages 7:0-15:11. Further standardisation data are being collected with a view to revising the norms for age 16:0-16:11 as soon as possible.

The distributions of raw scores for memory span did not permit calculation of standardised scores, because kurtosis exceeded acceptable limits. Given the nature of this particular measure (i.e. memory span) these statistical findings are entirely to be expected and the overall psychometric integrity of Recall is not affected. Consequently, comparative results for this measure is provided instead, as explained in Section 1.2.2.

Table 3. Means and standard deviations for the two derived measures in Recall.

Age	**Working Memory Composite**		**Working Memory Processing Speed**	
	Mean	*SD*	*Mean*	*SD*
7:0 - 7:11	20.13	8.49	0.85	0.17
8.0 - 8.11	21.94	8.88	0.78	0.15
9:0 - 9:11	26.23	9.47	0.74	0.13
10:0 - 10:11	28.78	9.63	0.70	0.12
11:0 - 11:11	30.58	10.07	0.68	0.11
12:0 - 12:11	31.66	9.97	0.64	0.13
13:0 - 13:11	32.14	11.00	0.65	0.13
14:0 - 14:11	33.29	12.10	0.64	0.12
15:0 - 15:11	34.88	10.71	0.64	0.13
16:0 - 16:11	31.66	10.32	0.65	0.14

1.4 Validity of Recall

1.4.1 What is validity?

Validation of a psychological or educational test is not the same thing as the psychometric standardisation of a test, nor should it be confused with the reliability of a test: 'reliability' generally refers to the extent to which a test can be expected to give the same results when administered on different occasions or by a different administrator, or the extent to which the components of a test give consistent results (see Section 1.5.1). 'Validity' is a measure of the extent to which the test measures what it is supposed to measure (e.g. reading or spelling ability). Validity is usually established by comparing the test with some independent criterion or with a recognised test of the same ability. Inevitably, this raises the issue of what is the 'gold standard' – i.e. which is the 'best' measure of any given ability against which all other should be compared? Professional opinions differ as to the merits of various tests and consequently there are no generally agreed 'gold standards' for assessing reading, spelling and writing. Hence the conventional method of establishing test validity is to show that a new test produces results that agree reasonably closely with well-established test(s) of the same ability.

External validity of Recall was explored in three ways. Firstly, analyses were used to examine the relationships between performance on Recall and children's attainment in school. Secondly, the Recall profiles of children with special educational needs were examined. Finally, convergent validity was explored by examining the relationships between scores on Recall and scores on the Working Memory Rating Scale (Alloway, Gathercole & Kirkwood, 2008).

1.4.2 Children's school attainment

Several previous studies using measures of the multiple-component model of working memory have demonstrated close relationships between working memory and children's attainment in school (e.g. Gathercole et al., 2003; Gathercole et al., 2004; St Clair-Thompson & Gathercole, 2006). These have typically used formal assessments of children's progress on the National Curriculum. However, policy changes have recently resulted in schools carrying out less testing. Instead, teachers are expected to record each child's progress each academic term, measured by tasks and tests that are administered informally. Teacher ratings of the National Curriculum levels of 337 children (166 males and 171 females) from the standardisation sample were therefore obtained from the schools at the time of testing. Table 5 shows the correlations between the three Recall subtests and children's National Curriculum levels at aged 7, 8, 9, 10 and 11. Consistent with previous research using other working memory tasks, there were statistically significant correlations between scores on Recall and children's scholastic attainment (see Table 4).

1.4.3 Children with Special Educational Needs

Research has suggested that performance on working memory measures can be used to accurately identify children who are likely to require special educational provision. For example, Gathercole and Pickering (2001) compared the working memory profiles of children with special educational needs to the profiles of children without special educational needs. Children with special educational needs performed significantly more poorly on measures of the central executive, and one measure of the visuo-spatial sketchpad.

Two schools who participated in the 2012 standardisation were therefore asked to supply the names of children recognised by the school or local education authority as having special educational needs. There were 37 such children. Their performance on Recall, compared to 46 age-matched children from the same schools, is shown in Figure 1. Consistent with previous studies children with special educational needs performed significantly poorer than children without special educational needs on the pattern recall task, $F(1, 92) = 12.23$, $p = 0.001$, and the counting recall task, $F(1,83) = 11.27$ $p = 0.001$. The difference between the two groups was not statistically significant for word recall, $F(1,91) = 1.10$, $p = 0.30$.

Table 4. Correlations between Recall scores and National Curriculum Levels.

		Reading	Writing	Mathematics
Aged 7 (N=50)	Word recall	.45**	.37*	.41**
	Pattern recall	.36*	.29*	.34*
	Counting recall	.52*	.42**	.52**
Aged 8 (N=69)	Word recall	.45**	.35**	.46**
	Pattern recall	.45**	.41**	.43**
	Counting recall	.58**	.50**	.56**
Aged 9 (N=71)	Word recall	.45**	.43**	.31**
	Pattern recall	.35**	.44**	.37**
	Counting recall	.48**	.60**	.48**
Aged 10 (N=81)	Word recall	.50**	.58**	.45**
	Pattern recall	.46**	.52**	.47**
	Counting recall	.45**	.39**	.45**
Aged 11 (N=66)	Word recall	.48**	.50**	.40**
	Pattern recall	.46**	.55**	.41**
	Counting recall	.31*	.32*	.32**

*Note: ** correlation is significant at the p<0.01 level, * correlation is significant at the p>0.05 level (2 tailed test).*

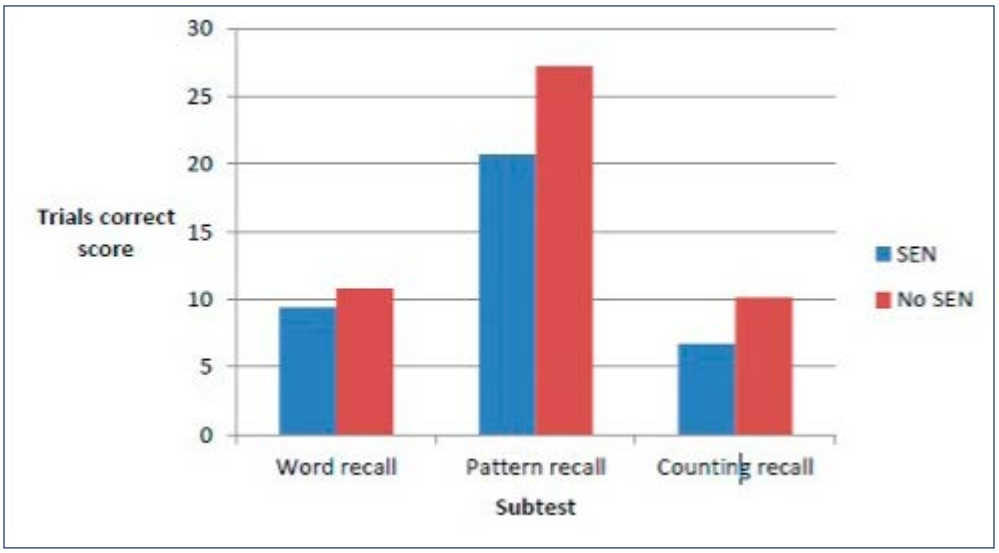


Figure 1. Profiles of children with and without special educational needs.

1.4.4 Scores on the Working Memory Rating Scale

As discussed earlier, there are several behaviours typically associated with a poor working memory. These include, for example, children losing their place in complex tasks with multiple steps, or requiring regular repetition of instructions in the classroom. Such behaviours can be

examined using the Working Memory Rating Scale (Alloway, Gathercole, & Kirkwood, 2008). Using this scale teachers are asked to rate how typical each behaviour is of a child using a four-point scale. Cognitive assessments of children's working memory have been found to be significantly related to teacher ratings on the Working Memory Rating Scale (e.g. Alloway, Gathercole, Kirkwood, & Elliott, 2009; St Clair-Thompson, 2011). Therefore, to further establish the validity of Recall the relationships between scores on each subtest and scores on the Working Memory Rating Scale were explored. Ratings were obtained for 51 children (27 males and 24 females) aged 7 years of age. The correlations are shown in Table 5.

Table 5. Correlations between scores on Recall and ratings on the Working Memory Rating Scale.

Working Memory Rating Scale score
Word recall -0.52**
Pattern recall -0.53**
Counting recall -0.46**

*Note: ** correlation is significant at the .01 level.*

Performance on each memory subtest was significantly negatively related to teacher ratings on the Working Memory Rating Scale. Negative correlations are expected because on the Recall subtests a higher score indicates a better working memory, whereas on the Working Memory Rating Scale a higher score indicates more problematic behaviours.

In summary, Recall has good validity. Correlations revealed that the scores on Recall are significantly related to children's academic performance across the childhood years. Children with special educational needs performed significantly more poorly than age-matched controls on the pattern recall and counting recall subtests. Finally, scores on Recall were significantly related to teacher ratings of children's behaviour on the Working Memory Rating Scale (Alloway et al., 2008). It can therefore be concluded that Recall is a valid assessment of children's working memory.

Note that Recall has been updated and restandardised since these studies examining validity were carried out.

1.5 Reliability of Recall

1.5.1 What is reliability?

'Reliability' generally refers to the extent to which a test can be expected to give the same results when administered on a different occasion, or by a different administrator, or to the extent which the components of a test give consistent results. Note that this is not the same as the validity of the test (see Section 1.4.1).

1.5.2 Reliability results

Test-retest reliability of Recall was assessed using a subgroup of children from the 2012 standardisation sample. A total of 119 children (62 males and 57 females) aged 7-9 years, and 45

children (22 males and 23 females) aged 13 years were given the three Recall core tests on two occasions. These were separated by an interval of 6 weeks. As test-retest reliability is expected to reduce over time this was a fairly stringent test of reliability. For each core test reliability was computed using the Pearson's product moment correlation coefficient. The resulting reliability estimates are shown in Table 6.

Each of the word recall, pattern recall, and counting recall subtests demonstrated good test-retest reliability. Although reliability was lower for Counting Recall in children aged 7-9 than children aged 13 years this finding could be attributed to some children finding it difficult to grasp this task, particularly at the first time of testing (as reflected in the large proportion of younger children that obtained low scores on this task). Previous research has also revealed similar test-retest reliability values for counting recall in children (e.g. Pickering & Gathercole, 2001).

Note that Recall has been restandardised since this reliability study was carried out.

Table 6. Reliability estimates on each subtest (test-retest after 6 weeks).

	Aged 7-9	Aged 13
Word recall	0.71	0.68
Pattern recall	0.69	0.77
Counting recall	0.49	0.76

1.6 Using Recall on Testwise

In order to access Recall through the Testwise platform, please follow this link: https://support.gl-assessment.co.uk/knowledge-base/platforms/testwise .

This will guide you through:

- Student management
- Sitting creation
- Taking the test
- Generating reports

2 Test administration

2.1 Details of the tests in Recall

2.1.1 Word Recall

This is a test of phonological loop functioning in which the child hears sequences of words through the computer speakers/headphones. They are then required to recall the words in the same order in which they were presented, selecting the target words from within a 3 x 3 matrix of nine words on the computer screen (for example, see Figure 2). All the words in this test (target words and distractors) are common single-syllable words between three and five letters in length, selected at random from a large data set, with built-in checks to avoid rhyming or alliterative pairs.

The test begins with a demonstration of what is required, followed by four **practice items**, two items in which there are two words to remember, and then two items in which there are three words to remember. During these practice items the child receives feedback to inform them whether they were correct or incorrect.

There are then a **maximum of six test items** at each list length (of 2–6 words), during which no feedback is given. Each list length is regarded as a **level**. If four items at any particular level are recalled correctly then the program jumps to the next level, omitting the remaining one or two items in that level (which are credited to the score). The task is discontinued when three or more trials at any level are recalled incorrectly.

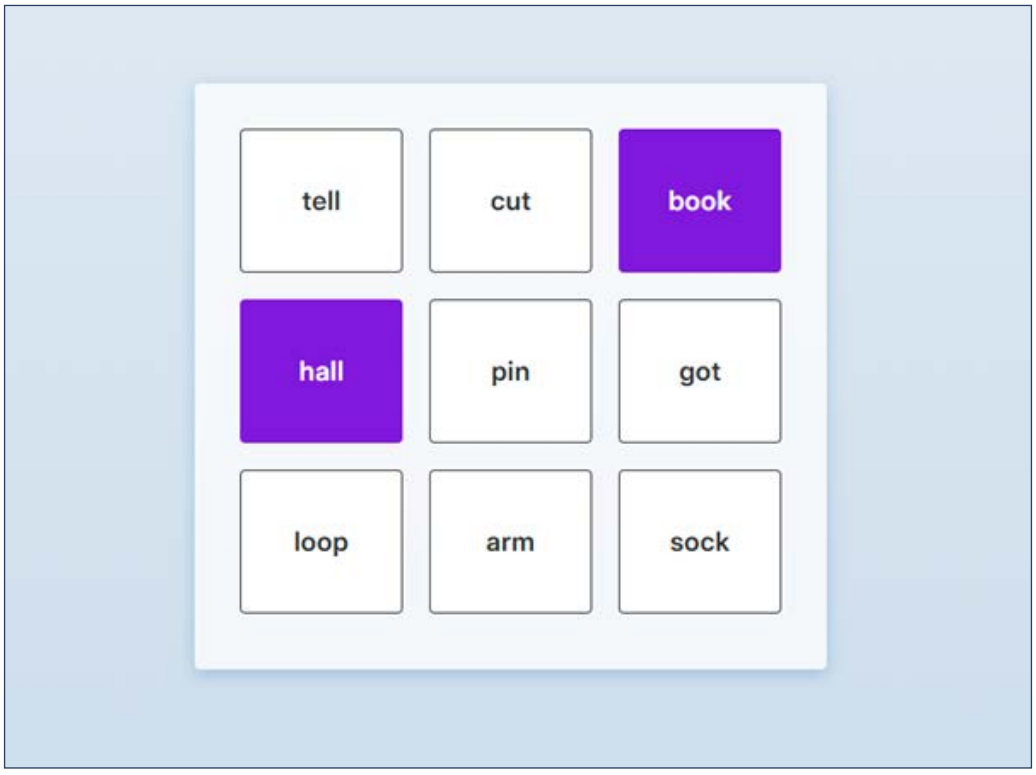


Figure 2. Example screen from the Word Recall test

2.1.2 Pattern Recall

This is a test of the functioning of the visuo-spatial sketchpad in which the child sees a matrix pattern of filled (dark) and unfilled (white) squares on the computer screen. When the pattern disappears they are presented with a blank matrix of all white squares and they are then required to recreate the pattern by using the computer mouse to click on the squares to be filled (for example, see Figure 3). All patterns used in this test are generated randomly.

The test begins with a demonstration of what is required, followed by four **practice items**, two items with a pattern of two filled squares on a matrix of sixteen squares (4 x 4), and then two with a pattern of three filled squares. During these practice items the child receives feedback to inform them whether they were correct or incorrect.

There are then a **maximum of six test items** at each pattern size, starting with two squares and increasing by one square per level up to a maximum of 12 squares, during which no feedback is given. Each pattern size is regarded as a **level**. If four items at any particular level are recalled correctly then the program jumps to the next level, omitting the remaining one or two items in that level (which are credited to the score). The task is discontinued when three or more trials at any level are recalled incorrectly.

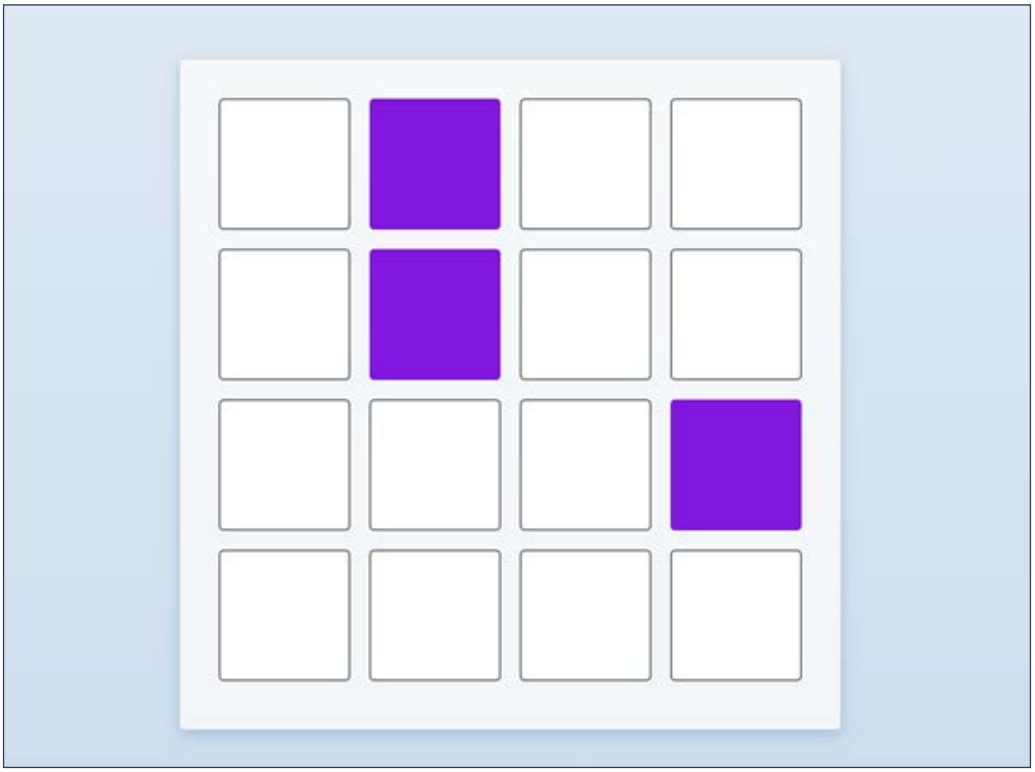


Figure 3 Example screen from the Pattern Recall test

2.1.3 Counting Recall

This is a test of central executive functioning that involves carrying out a sequence of between two and six independent counting tasks whilst simultaneously remembering the results of each count in the same order. In each count the child is presented with an array of different shapes, the numbers and locations of which are randomly generated, and is required to count the number

of orange circles, selecting the correct answer at the bottom of the screen (for example, see Figure 4). At the end of each sequence of arrays they are asked to recall the number of orange circles in each counting array, in the same order in which they were presented.

The test begins with a demonstration of what is required, followed by **four practice items**, two with two-count arrays, and then two with three-count arrays, with feedback given after each. There are then a **maximum of six test items at each sequence length** (from two to a maximum of six counts in the sequence), during which no feedback is given. Each sequence length is regarded as a **level**. If four items at any particular level are recalled correctly then the test jumps to the next level, omitting the remaining one or two items in that level (which are credited to the score). The task is discontinued when three or more trials at any level are recalled incorrectly.

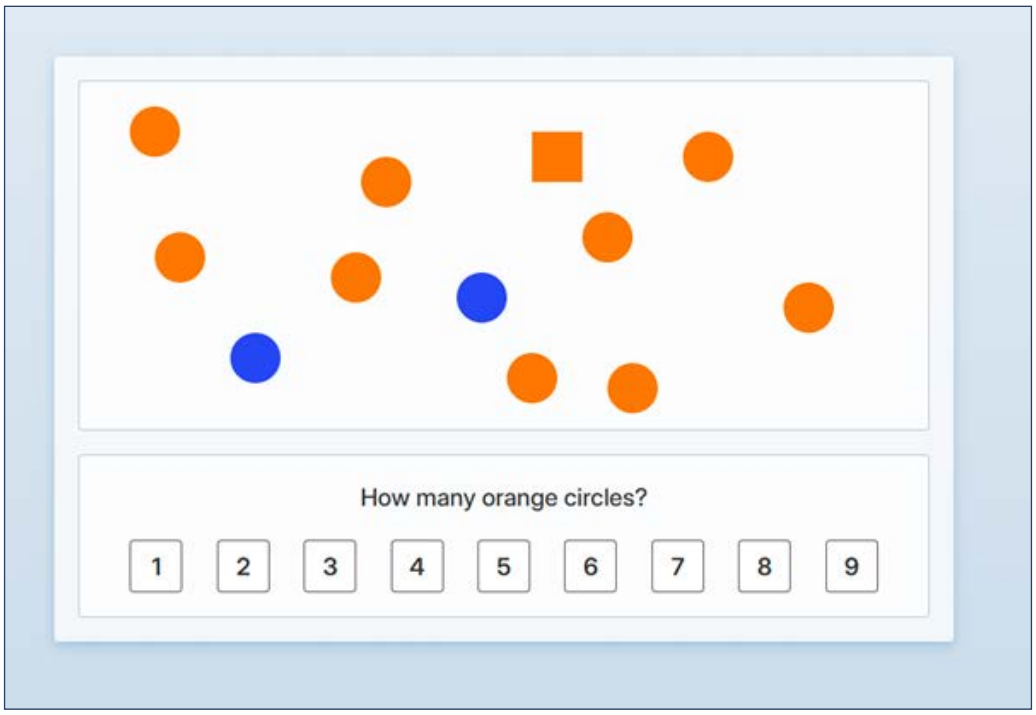


Figure 4. Example screen from the Counting Recall test

2.2 Administration guidelines

2.2.1 Trial run-through

Assessing students with Recall is straightforward but before the teacher or administrator attempts to test any student it is advisable first to run through the complete suite of tests to familiarise themselves thoroughly. To do this, create a demo student account in Testwise and add yourself to a sitting.

2.2.2 Testing environment and equipment

The ideal testing environment is one that is reasonably quiet, with minimal distractions. Ideally, this should be a separate room, but Recall has been designed to be robust for use in the ordinary classroom, provided visual and auditory distraction (both to the student being tested and to

other students in the class) have been minimised. To minimise auditory distraction, headphones are recommended. Inexpensive lightweight headphones of the type used for portable audio equipment will be adequate (but not the type that are inserted into the ear). Teacher or supervisor judgement is paramount in ensuring the appropriate testing environment.

If assessment is going to be carried out in an ordinary classroom in which there are other pupils, the computer and the student should be positioned in such a way that the student is not looking directly at the rest of the class, nor should the rest of the class easily be able to see the monitor screen. The best position for this is usually in the corner of the room. Students should not attempt the tests when other students are in a position in which they can become involved in the task or act as a distraction. It will be hard for other students to inhibit their reactions and their behaviour may influence the decisions of the student being tested.

The teacher or supervisor should check that the equipment being used for the assessment is functioning correctly. This includes checking (1) that the sound system (speakers or headphones) is audible (not too loud or too soft, and without interference), and (2) that the mouse is functioning correctly.

Please ensure that you have a reliable internet connection and that all student devices meet the minimum system requirements before beginning the test.

2.2.3 Student preparation

Before testing, student details for all pupils must be uploaded to Testwise. You will then be able to create a new sitting choosing the Recall product, set start and end dates, and select pupils. Once the sitting is saved, access codes can be exported in csv or pdf format.

Pupils will need to navigate to www.testwise.com/code and enter their assigned access code. They will be asked to confirm their identity, and then the test will load.

Written and audio instructions are provided, and each test commences with a demonstration of the task followed by a series of practice items. When the student has completed the practice items, the test phase begins.

The student should be sitting comfortably at a suitable level in front of the computer screen (not too high or low, in order for them to see the screen). It is not recommended that students attempt the tests standing up, as they are more likely to move about and alter the angle at which the screen is viewed – this can lead to failure to see everything that is happening on the monitor, and can also disrupt mouse control. The supervisor should check for reflections on the monitor from windows and lights that could impair the student's perception. To do this the supervisor should check by viewing the screen from the same position that the student will adopt.

If necessary, students should be shown how to indicate responses using the mouse or touchscreen, and when to respond (essentially when the tests will allow them to respond). This is particularly important when testing students with physical disabilities. As with any formal assessment, students should not be allowed to take the tests if they are unwell, as results are likely to be unreliable.

Most students will experience no difficulties in understanding what is required of them when taking the tests in Recall, enabling them to follow the practice tasks easily and progress to the test phase without special attention from the teacher or supervisor. However, it is important that the administrator ensures that students understand the nature of the tasks in Recall: that they are tests and not games, and they must work swiftly but thoughtfully and try their best at all times.

In the rare event that a student does not understand the instructions spoken by the computer the supervisor may re–express them in a more suitable manner. Explaining and re-expressing the task requirements to the student may continue into the demonstration and practice stages of each test. This is particularly useful for any student who is experiencing problems in understanding the true nature of the task. It is often easier for the student to comprehend the task requirements by experience of the practice stages, than by more abstract oral explanation. Once the test items commence there should be no further aid given to the student. The three tests in Recall can be completed in any order.

2.2.4 Supervision

It is usually not necessary for students to be closely supervised while attempting the tests, unless the teacher or administrator has a particular reason to do so. Recall is specifically designed for group testing and to require minimal input from the teacher or administrator. Note that if the results are being used to apply for exam access arrangements, JCQ regulations require the specialist who signs the JCQ forms to supervise the assessment.

Children with special educational needs may require additional support during assessment. If children struggle with word reading they may find it particularly difficult to complete the word recall test, which relies upon children being able to remember words, but also read the target words and distractor items. Support can therefore be offered so that children can recall the words out loud and a teacher or support worker can click on these words on the screen to provide a response. Using this method the scores will still reflect a child's working memory, but will not be influenced by reading ability.

The tests in Recall have been designed to be interesting and stimulating for students in this age group and the vast majority of students are highly motivated to do their best. Once the teacher is satisfied that the student understands the requirements of a test, has completed the practice items and has moved on to the test items, the teacher may leave the student to complete that test.

Where the teacher suspects that a student may not be well motivated to complete the test, or may be easily distracted, or may be performing deliberately below their capabilities, closer supervision will be necessary. Disaffected students may display non-compliance by clicking on answers at random, rather than thinking about the tasks and selecting answers after proper consideration. Such students, or those with very low ability, may need close supervision in order to provide encouragement and ensure they remain on task.

In order for the assessment to be 'fair' (i.e. to give a reasonably accurate representation of the student's abilities) it is essential to ensure that during the test:

- the student is paying attention, is 'on task', is not distracted and trying their best
- the student does not become unduly fatigued

- there is no teaching or helping with the task during the test items (whether from the supervisor or other students)
- that feedback from the supervisor is minimised and encouragement consistent (see further comments below).

2.2.5 Giving encouragement, prompts and feedback

As much as possible, the supervisor should avoid giving specific feedback to students during a test, because this may influence their behaviour in an undesirable fashion. This is good practice in any testing situation. There is a risk of feedback differentially affecting students, so that some are encouraged and others discouraged. Nevertheless, some students (particularly younger students or students with special educational needs) will try to elicit feedback from the supervisor about their performance. This may take the form of both verbal and non-verbal behaviours. For example, the student may ask directly if they were correct. Many students will look for the supervisor's facial and bodily reactions to their responses. Some students may even try to evaluate the supervisor's reaction by observing the supervisor's reflection in the monitor screen. For these reasons it is usually preferable that if the supervisor is going to be near the student to observe the assessment they should sit to the side and slightly behind the student to minimise any feedback to the student which may bias the results.

Rather than specific feedback, general encouragement should be given to the student. This encouragement should be referenced to task completion rather than task accuracy and ideally should be delivered equitably to all students. However, it is inevitable that some students will require more encouragement than others, and where this is the case the teacher should be mindful of the possibility of influencing results unduly. Differential encouragement between students is likely to have an influence on the results obtained, and therefore should be avoided where possible. Some key phrases and general incentive prompts which may be used to aid the administration of the tests include: "well done"; "you were good at that, now try the next one"; "you will like this game"; "now concentrate on this"; "try hard"; "listen very carefully"; "have a go at these ones"; "have a try"; "just do your best".

2.2.6 Timing of the assessment

It usually takes about 20–30 minutes for most children to complete the three tests in Recall, although children who take a long time to respond to each trial, or more able children who progress to the most difficult levels, may take slightly longer. However, it is recommended that no time limit should be imposed upon children, who should be instructed to complete the tasks in their own time.

Each test is administered using a span procedure. This means that at the beginning of each test there are only a few items to remember, but the number of items then increases over successive trials. Testing is automatically terminated when a child continues to respond incorrectly, indicating that the number of items to be remembered has exceeded their memory span. This minimises the amount of time it takes to complete the assessments.

The time required to complete each test is also minimised by using a progression rule. A trial is one sequence of items or one pattern that is presented for recall. Each test is organised in blocks

of six trials which have the same level of difficulty (i.e. the same number of items to remember). If a child correctly recalls four trials in any one block they automatically progress to the next block and full credit is given for the omitted trials. When three or more errors are made within a single block testing is automatically discontinued.

2.3 Assessing students outside the 7–16 age range

2.3.1 Assessing students under 7:0

It is standard practice that normative tests are not generally recommended for use outside the age range for which they have been designed and standardised. Any test, such as Recall, which meets basic psychometric criteria must be standardised on a given population and this will determine the range of applicability of the test (see Section 1.3 for explanation of the standardisation process). Tests appropriate to the students' chronological age should be used wherever possible, to avoid the dangers of inappropriate decisions being made – e.g. that a student is 'at risk' (or not 'at risk') or requires intervention (or no intervention) when the evidence for this may be unsound.

If the student being assessed is younger than age 7:0, then Recall will use the norms for the age 7:0 when analysing results, and this will almost certainly lead to an underestimation of their performance as chronological age generally has a major impact on performance in childhood. However, it is also important to be aware that the tests in Recall were not designed for children under the age of 7; such children may find the tests too difficult, scoring at, or close to, what is called the floor of the test (i.e. the minimum raw score obtainable on a test; in Recall this is zero). On the word recall test, more than 10% of children under 7:0 would be expected to score at floor level. On the pattern recall test, more than 4% of children under 7:0 would be expected to score at floor level. On the counting recall test, more than 21% of children under 7:0 would be expected to score at floor level.

From these figures it can be seen that Recall loses much of its discriminatory power at lower skill levels when used with children under 7:0; this is most apparent on the counting recall test (on which the test does not discriminate below SS 88), somewhat less marked on the word recall test (on which the test does not discriminate below SS 81, and least of all on the pattern recall test (on which the test does not discriminate below SS 73).

On the other hand, if a child under 7 is believed to be ahead of their peers in cognitive development, Recall may be useful in revealing how advanced their development is. In such circumstances **age equivalents** would be the preferred form of scores for the teacher or administrator to use, rather than standard scores or centile scores, and results should always be interpreted with care. For further information about age equivalents, see Section 3.2.4).

Consequently, use of Recall with students under the age of 7 is not recommended except under special circumstances as outlined above.

2.3.2 Assessing students older than 16:11

Recall was designed for use with students aged up to 16 years 11 months and use with students older than this can create problems when interpreting results. However, provided assessors are aware of the issues involved and results are interpreted with care, Recall can be used with students older than 16:11.

Be aware that some regulations (e.g. JCQ regulations for exam access arrangements) insist that students being assessed are within the normed age range for the test.

Just as the use of Recall with children younger than age 7 can confront the assessor with test floor issues, Recall with students older than age 16:11 may occasionally create test ceiling issues. Strictly speaking, the ceiling of a test is the maximum raw score obtainable, and that is the meaning used here. Sometimes, however, the term is applied to the upper age limit for which the test has been normed, because over this age limit the standard score norms will not be valid.

On the word recall test, more than 3% of students over age 16:11 would be expected to score at ceiling level (raw score 30). On the pattern recall test, more than 1% of students over age 16:11 would be expected to score at ceiling level (raw score 66). On the counting recall test, more than 5% of students over age 16:11 would be expected to score at ceiling level (raw score 30).

While the proportions of students over age 16:11 scoring at ceiling level on Recall are generally smaller than the proportions of students under age 7:0 scoring at floor level, it is still apparent that some discriminatory power is lost when Recall used with more skilful students over age 16:11. Above this age the word recall test does not discriminate above SS 128, the pattern recall test does not discriminate above SS 139, and the counting recall test does not discriminate above SS 124. However, this is probably less of a problem than younger students scoring at the floor of the test because the principal use of Recall is likely to be in identifying students who have working memory difficulties. If the focus of interest is in those students whose working memory is well below that of their peers, the lack of discriminatory power with older students who have superior memory skills is not a serious concern.

If the student is older than 16:11 then the program will use the norms for age 16:11 when analysing scores. For this reason, when assessing students over age 16:11, **age equivalents** are the preferred form of scores for the teacher or administrator to use. For further information about age equivalents, see Section 3.2.4).

2.4 Retesting

It is often desirable to retest a student after the initial assessment, either to monitor progress or because the first assessment was unreliable for some reason (e.g. because the student was unwell, unmotivated or misunderstood the requirements). As explained in Section 1.1.6, due to practice effects, retesting using conventional tests can be problematic and minimum intervals between testing sessions (e.g. 6 or 12 months) are stipulated to try to counter these effects, or parallel forms of the test are used.

Since item content in Recall is generated at random, practice effects are minimised and consequently there are no stipulated intervals between testing and retesting with Recall, nor any restrictions on the number of times students may be retested. This is particularly useful when needing to evaluate the impact of intervention over time.

2.5 Assessing students who have limited English

Assessment of any student who has limited proficiency in spoken or written English can be problematic, but there is evidence that Recall is much better than many conventional methods of assessment because of its strongly visual format and minimal reliance on spoken instructions. In order to tackle the Word Recall test the student will need to have had sufficient familiarity with simple English words in their written form, and to tackle the test of Counting Recall the student will need to know the digits 1–9 in written form. The practice items enable most students, even those with very little English, to understand the tasks, and where there is uncertainty a teacher or assistant who speaks the student's mother tongue can help with explaining instructions. It can be seen that these are pretty basic requirements that are unlikely to be a problem for most students for whom English is an additional language (EAL).

For a discussion of working memory in relation to learning a second language, see Juffs and Harrington (2011).

3 Understanding results

3.1 Types of report

An individual report is available for each student, and it is also possible to export an Excel file containing results for a group of students.

The **Individual Student Report** occupies a single A4 page and presents all the results for a given student on a single assessment (for example, see Figure 5). A box is provided at the bottom of the page in which the Assessor can add comments. All results are shown in tabular format, with standard scores also being shown in graphical format as a bar chart. To aid speedy identification of areas of difficulty, the bars on the chart are coloured blue if the standard score is 85 or above (i.e. within the normal range or better), and yellow if below 85 (i.e. below the normal range, indicating that the result is a matter of concern).

The **Excel Report** contains all Recall results for a group of students in a single Excel file. It provides all the same data as the individual reports but enables teachers, administrators or SENCOs to see an overview of a group's results and identify pupils whose results indicate potential difficulties.

3.2 Types of scores

All raw scores on Recall are saved when the test is submitted. The data saved also includes the date and time the test was completed, as well as the registered details of the student. If a test has been abandoned before completion, then no results will be saved for that test. Reports are calculated at the time they are generated. This is important, for example, where errors have been made in entering the student's date of birth, in which case the wrong norms may have been used. Therefore if any mistakes of this nature were made then it is important to recalculate the results by generating new reports after any corrections have been made.

Recall then refers to the standardised norms in order to convert raw scores to the following three types of score.

- Standard scores (SAS) and confidence band
- Centile scores
- Age equivalents

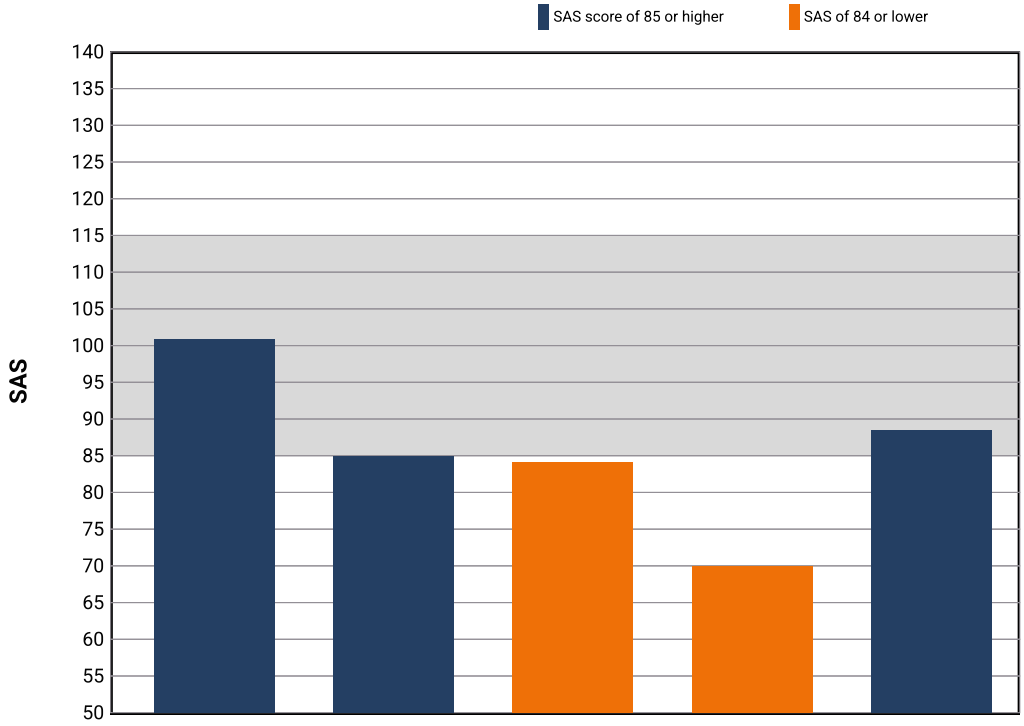
The first of these is shown in graphical (bar chart) format as well as numerical format, while the remaining two are shown only in numerical format. These different types of score formats are explained in the following sections.

Recall Individual Student Report

GL Education A Renaissance Company Recall

Name: Rachel Smith	**Date of Birth:** 03/12/2017
School: GL School	**Group:**
Date of Test: 03/10/2025	**Age at Test:** 7:10

Results Profile



	Tests				Working Memory
	Word Recall	Pattern Recall	Counting Recall		
	Phonological Loop	Visuo-Spatial Sketch Pad	Executive Function	Processing Speed	Composite
SAS	101	85	84	<70	88
Confidence Band	87-115	73-97	72-96	<70-81	76-100
Percentile Score	54	16	14	1	22
Raw Score	11	9	2	1.15	17
Memory Span	3	3	2		
Age Equivalent	09:00-09:05	05:00-05:02	05:06-05:11	<05:00	07:03-07:05

Assessor's comments

Figure 5. Example Single Assessment Report.

3.2.1 Standard age scores (SAS)

Standard age scores (SAS) are provided from age 7:0 to 16:11. Standard scores have a mean (average) 100 and a standard deviation of 15 (see Figure 8). The standard deviation is a statistical measure of the average variability of scores in a distribution. Approximately two-thirds of the population will have scores that fall between plus or minus one standard deviation of the mean (i.e. score range 85 – 115, which is the area shaded blue on the graph in Figure 8). In some scoring systems the range 85 – 115 is regarded as the 'normal' or 'average' range, while other systems treat 90 – 110 as the 'normal' or 'average' range; in the latter case, 50% of the population will fall into the average band. The more extreme the score the fewer individuals are found in that category, so that only about 2% of the population have very low scores (less than 70) and about 2% have very high scores (130+). This distribution of scores is a characteristic of many human attributes (height, weight, strength, sociability, etc.), i.e. most people tend to cluster around a central point and as one approaches the extremes (known as the 'tails' of the distribution) fewer people are found.

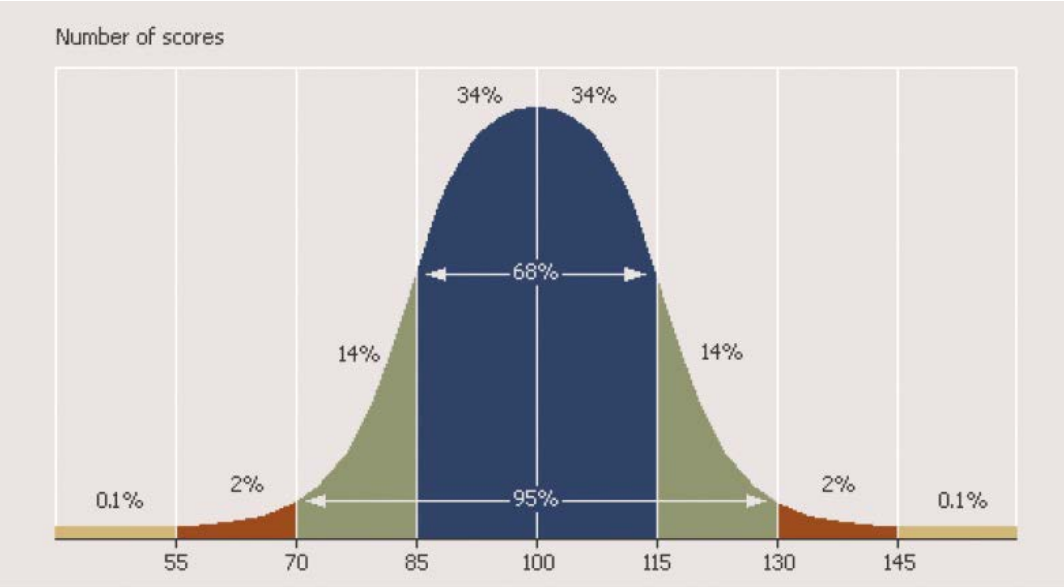


Figure 8. Distribution of Recall scores on a normal curve (the figures along the bottom of the diagram correspond to standard scores)

3.2.2 Confidence intervals

When reporting a standard score, it is good practice also to report the confidence band (or interval) associated with that score. The reason for this is that all psychological and educational tests scores give only estimates of ability, based on a sample of behaviour at a given point in time. If you were to assess a student on several occasions you would not expect them to obtain exactly the same score each time – there would be a spread of scores and somewhere within that spread we would expect the (hypothetical) true score to lie. The amount of spread or variation of actual scores obtained by an individual is dependent on the reliability of the test. The confidence interval is the zone around the standard score in which we are reasonably confident the true score lies. Different confidence intervals may be set: for Recall we have set a confidence level

of 90%, which means that there is a 90% probability that the true standard score lies within the stated confidence interval. Put another way, if the student was retested 100 times, on 90 out of 100 occasions the score would lie within the stated confidence interval.

Confidence intervals are calculated on the basis of the Standard Error of Measurement (SEM) of a test which, in turn, is determined by the reliability of the test and the standard deviation of test scores (see Section 1.5).

3.2.3 Centile scores

Centile scores are provided from 7:0 to 16:11. Centile scores (sometimes referred to as 'percentile' scores) represent the student's performance in comparison with the population norms in centile units which range (roughly) from 1 to 99. A centile score of 63, for example, means that the students' score lay at the point where 63% of the population scored less, and 37% scored more. A centile score of 50 indicates that the student's score lay exactly on the median (middle point) of the distribution, with half the age group scoring higher and half lower. As will be obvious from Figure 8, centile scores have a strict relationship with standard scores as shown in Table 7.

Table 7. Relationship between standard scores and centile scores.

Standard score	70	80	85	90	100	110	115	120	130
Centile score	2	9	16	25	50	75	84	91	98

3.2.4 Age equivalents

Age equivalents are provided for the age range 5:0 to 16:5 or 16:11, depending on the test (over this age, age equivalents become meaningless). Age equivalents may be defined as the average chronological age of students who would be expected to achieve a given raw score on the test. Age equivalents are another way of expressing how a given student is performing in relation to their peers.

Note that because of the way that age equivalents are calculated they are not as precise as standard scores or centile scores; age equivalents should be regarded as approximations and hence are often given in bands. Age equivalents should be used with caution and only in cases where standard scores or centile scores would be inappropriate or unhelpful. It is embarrassing and demotivating for a teenager or adult to be told (for example) that they are performing at the age of a 7-year-old. However, some teachers working in special education prefer to use age equivalents rather than centile scores, because age equivalents enable them to conceptualise the ability level of the student they are teaching, and so pitch the work at the correct level. Also, when circumstances dictate the use of Recall for assessing a student younger than 7:0 or older than 16:11, age equivalents can prove useful (see Section 2.3 for further information on this).

3.2.5 Raw scores

Raw score are the actual scores obtained by the student on each test. For all except Working Memory Processing Speed the raw scores represent the number of correct items on the test (for Working Memory Composite, scores have been weighted to reflect the different numbers of items in each test). For Working Memory Processing Speed the raw score represents the average time in seconds per item counted. For most purposes, raw scores are not particularly useful or interesting, but they may be relevant for some researchers. Note that two students can obtain the same raw score on a test but have different standard or centile scores if their chronological ages are different.

3.3 Memory span

The tests in Recall are constructed in levels, each level representing an increase in the number of items of information that have to be held in working memory. The maximum number of items of information that an individual can hold in working memory is called their 'memory span', and for each of the three core tests in Recall this figure is shown on the report. The memory span range is fairly small: for the age range used in Recall (7:0 – 16:11) the span ranges are 2–6 for word recall and counting recall, and 2–12 for pattern recall. For this reason, the data are not suitable for standardisation in the conventional sense, because the basic psychometric principles of normality of distribution, including kurtosis being below acceptable limits, are not met (for further explanation of the reasons for this, see Section 1.3.3).

Memory span is dependent on the nature of the information being processed. For example, for an 8-year-old, a memory span of five is above average for words but only average for patterns. This is partly because words require more memory storage space than simple patterns, and partly because the child learns to process visual information before being faced with the complexities of relating spoken words to their printed forms. Memory span also tends to increase with age. For example, a memory span of seven for patterns is above average at age 8 but below average at age 12.

3.4 Interpreting scores

If a child has a poor or below average working memory it is likely that they will struggle in the school classroom and be at risk of poor educational attainment. In such cases it may be appropriate to consider a full and detailed assessment of a child's working memory and related skills. In such cases it is also necessary to explore interventions to reduce the chance of children failing on learning activities as a result of a poor working memory. If a child has an average working memory they may also benefit from some of the interventions. It may be possible for them to improve their working memory to an above average or good level. If a child already has an above average or good working memory, they will be well equipped to perform well in the school classroom and achieve good levels of scholastic attainment. Details about teaching strategies and interventions that can be used for children with a poor working memory are discussed in detail in the next chapter.

4 Helping children with a poor working memory

4.1 Poor working memory in the classroom

As discussed in the earlier sections, working memory plays an important role in children's learning. Children with a poor working memory therefore make frequent errors in learning activities. They may struggle to follow instructions or complete multi-step tasks. They may have difficulty remembering, so need to be reminded about instructions or tasks at hand. They may make place-keeping errors, such as missing out letters or words in a sentence, or writing the same word twice. This might mean that they take longer than other children to complete learning tasks, and ultimately, they may struggle to progress in key domains like reading and writing.

It is important for teachers to recognise working memory difficulties in the classroom. Children with a poor working memory are sometimes described as having attentional problems or being likely to engage in "mind-wandering" (e.g. Kane, Brown, McVay & Silvia et al., 2007). This has been referred to as 'zoning out' and is common is situations in which working memory is overloaded and therefore it is not possible to keep the information needed in mind. Children therefore fail to remember crucial information, and so they shift attention away from the task. However, the behaviours of children with a poor working memory are otherwise unlike those of children who may be diagnosed with attentional difficulties. For example, they may sometimes be able to concentrate even when struggling with a learning task, and may not experience other challenges such as emotional or peer relationship difficulties.

Poor working memory is therefore difficult to recognise in the school classroom. However, Recall provides a reliable, valid, and efficient method for identifying children with a poor working memory. It can be administered in group settings, and no teacher or researcher input is required.

After identifying that a child has a poor working memory, steps can be taken to minimise the chance of a child failing on learning activities because of a poor working memory. There are two main approaches to intervention. These are discussed in detail in the following sections.

4.2 Teaching strategies

The first approach to minimising the difficulties experienced by children with a poor working memory is to reduce the working memory demands or cognitive load of classroom activities. Teachers should provide clear instructions and repeat these or make these available on a worksheet or on the screen at the front of the class, so that they do not need to be remembered. These should be as simple as possible. When a task has multiple elements, teachers should consider whether the amount of information could be reduced whilst maintaining the core content. Design elements should also be considered. Complex fonts, images, and diagrams can mean that there is unnecessary information to process.

When working on tasks, routines that are practised frequently tend to become automatic. This then frees up working memory resources to focus on a task at hand. So, routines for instructions, writing, or mathematical calculations should be modelled and well-practised before children undertake more complex or independent tasks.

Working memory demands can also be reduced by using memory aids. These could be classroom displays, numbers lines, formulas, or readily available lists of commonly misspelt words. Teachers can also teach children to use memory-supporting strategies. For example, if verbal material (e.g., words or numbers) must be remembered then rehearsal is a useful strategy. Information can be rehearsed aloud or internally. Note-taking can also relieve the working memory demands of a task. If a child often forgets information they can be encouraged to write down task instructions or notes that will help them complete a task. Arming a child with such strategies will promote their development as an independent learner.

Other important strategies centre on the time it may take children with a poor working memory to complete learning tasks. As mentioned earlier, children who frequently forget information, and therefore make errors in classroom tasks are likely to need longer than their classmates to complete activities. Classroom accommodations can include providing extra time (ideally without drawing attention to this), offering additional support or tutoring, and encouraging consistent practice. Adjustments can also be made in terms of examination requirements of course, and Recall can assist with this.

4.3 Working memory training

The second approach to alleviating the difficulties that arise from a poor working memory is to improve working memory directly. Interventions have included approaches as diverse as mindfulness or meditation training, neurofeedback, physical exercise, and long-term musical training. However, one approach that has captured lots of attention is cognitive training. Cognitive training consists of exercises involving the repetition of demanding tasks. There are many commercial products that have been developed for this purpose. Focusing specifically on working memory, one example is the Cogmed working memory training program. This involves several verbal and visuo-spatial memory span tasks that have been embedded within video games. It is an adaptive program, in that trial-by-trial performance determines how much information a participant is required to remember. Participants are expected to engage in intensive training, completing sessions daily. Early scientific studies suggested great potential for working memory training, although more recent research has questioned its effectiveness (Redick, 2019).

The key consideration for working memory training is whether improvement on practiced working memory tasks leads to improvements on other tasks. This is known as transfer. Scientific studies have examined whether working memory training improves performance on tasks that are like the practised tasks (e.g., other tasks of working memory). This is sometimes referred to as near-transfer. Studies have also examined whether it improves performance on tasks that differ from the practised tasks (e.g., performance on maths or reading tasks, or behavioural outcomes). This is referred to as far-transfer. Several reviews have now concluded that there are near-transfer effects of working memory, but that there is limited evidence for far-transfer, limiting the impact of working memory training in educational settings (e.g., Melby-Lervag & Hulme, 2013; Rapport et al., 2013; Redick et al., 2015).

Research has also examined the impact of memory strategy training. This differs from the cognitive training approach described above because here trainees learn specific strategies to help them encode, store, and retrieve information. Such strategies include rehearsal, chunking, mental imagery, or story formation. There is some evidence that strategy training can improve working memory and have some transfer effects. For example, St Clair-Thompson et al., (2010) asked children to use a program called Memory Booster. They found significant improvements in children's working memory after using Memory Booster for a period of 6-8 weeks. Children also showed significant improvements on tasks of mental arithmetic and the ability to follow instructions in the school classroom. However, evidence suggests that memory strategies are often context specific and thus that transfer may be somewhat limited (see also Shipstead, Redick, & Engle, 2012).

Although the potential for working memory training will continue to be researched in the future, due to the limited evidence of transfer effects the best way to support working memory in the classroom is to structure the learning environment appropriately and adapt teaching practices as considered in the previous section.

5 Case studies

The following five case studies illustrate how the results of Recall may be interpreted. It should be recognised that in each case conclusions can only be tentative because the information from a 20-minute screening program is inevitably limited. Nevertheless, in most cases the information provided by Recall can help a teacher or parent understand a child's difficulties. This will allow them to move forward and to consider strategies that could be used to minimise the chance of the child failing on learning activities because of inadequate working memory resources.

5.1 Rachel (7 years 10 months)

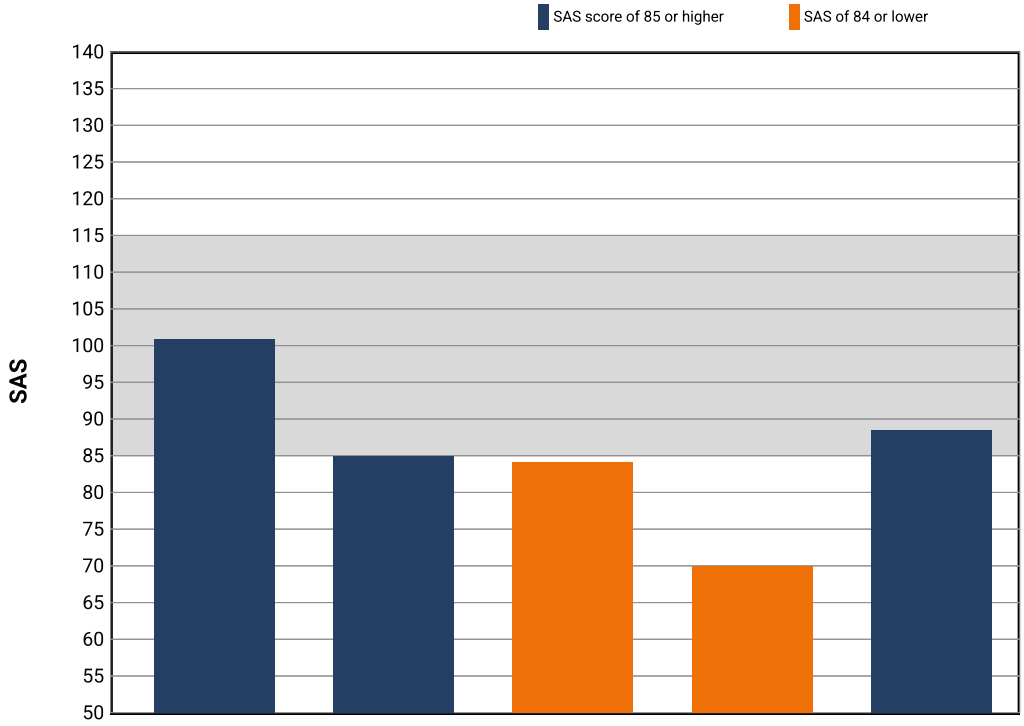
Rachel's performance on Recall is shown in Figure 9. Overall, her results indicate she has below average working memory. An inspection of Rachel's scores indicates satisfactory performance on the word recall test, so she has no problems with remembering and recalling sequences of verbal information. However, Rachel's performance on the pattern recall task was below average, and her performance on the counting recall task was poor. This profile of scores is typical of some children with special educational needs, who perform more poorly than age matched controls on measures of the visuo-spatial sketchpad and central executive components of working memory. Of particular concern is Rachel's score on the counting recall task, which indicates she has severe difficulties with coping with simultaneous processing and storage demands. Her working memory processing speed is also well below average. These results indicate that Rachel is likely to make slow progress with acquiring knowledge and skills in areas such as literacy and mathematics, and without appropriate intervention is at risk of poor educational attainment. Rachel's results lead to several recommendations. Firstly, it is important that Rachel's teachers recognise that she has a poor working memory, and that her difficulties are not a result of other problems such as inattentiveness. Teachers can then try to evaluate the working memory load of classroom activities, and where possible reduce working memory demands (see Section 4.2). Intervention work should focus on simultaneous processing and storage demands. For example, Rachel will benefit from some training related to using visual aids, so that they can be effectively used to reduce the amount of information that needs to be remembered during on-going processing tasks. It would also be useful for a teacher or teaching assistant to work closely with Rachel to encourage her to develop strategies for dealing with her poor working memory, including note taking. It is also important to reiterate to Rachel that it is OK to ask for help when it is needed. Finally, Rachel's teachers may also want to suggest some kind of working memory training (see Section 4.3). This would be likely to lead to some improvements in working memory. Although at the moment the long-term consequences of working memory training are unknown, training should at least improve Rachel's confidence and beliefs in her ability to deal with complex processing and storage tasks.

Recall Individual Student Report

GL Education A Renaissance Company Recall

Name: Rachel Smith	**Date of Birth:** 03/12/2017
School: GL School	**Group:**
Date of Test: 03/10/2025	**Age at Test:** 7:10

Results Profile



	Tests				Working Memory
	Word Recall	Pattern Recall	Counting Recall		
	Phonological Loop	Visuo-Spatial Sketch Pad	Executive Function	Processing Speed	Composite
SAS	101	85	84	<70	88
Confidence Band	87-115	73-97	72-96	<70-81	76-100
Percentile Score	54	16	14	1	22
Raw Score	11	9	2	1.15	17
Memory Span	3	3	2		
Age Equivalent	09:00-09:05	05:00-05:02	05:06-05:11	<05:00	07:03-07:05

Assessor's comments

Figure 9. Recall results for Rachel (age 7:10).

5.2 Emma (11 years 11 months)

Emma's performance on Recall is shown in Figure 10. Emma's results indicate that she has good working memory. An inspection of her scores reveals that she achieved above average scores for the pattern recall and counting recall task. Her performance was not as high on the word recall task, but still within the average range. These results suggest that Emma is unlikely to have problems with simultaneous processing and storage of information or with remembering information in the visuo-spatial domain. It is therefore unlikely that Emma will be recognised as having special educational needs, and in general she should perform well in the school classroom.

It may be useful to acknowledge Emma's poorer score on the word recall task. Firstly, it would be useful for teachers to check that a low score was not a result of Emma struggling to read the words during the recall phase of this task. As detailed in Section 2.2.4, word recall relies upon children being able to remember words but also read the target words and distractor items. Emma could therefore be asked to complete the task again, with a teacher or support worker asking Emma to repeat the words she can remember and then clicking on these words on the screen to provide a response for her. Using this method the scores will reflect Emma's working memory, in particular her phonological loop, but will not be influenced by reading ability. If Emma's score no longer indicates poor performance on this task, then teachers should be aware that Emma may have problems with single-word reading and needs some practice to improve this skill. If Emma's score is still indicative of poorer word recall, this will lead to several recommendations.

Poor performance on the word recall task, but not on the other tasks in Recall, usually indicates a specific problem with the immediate serial recall of verbal information. Firstly, it is important for teachers to recognise this difficulty with remembering verbal information. A student affected in this way may struggle with remembering instructions for a task or remembering sentences to write down. Teachers should therefore break down instructions into separate steps and regularly repeat important information. They should also use memory aids, for example, note task instructions on the class whiteboard. Training which includes appropriate strategies for remembering verbal information (e.g. rehearsing information to be remembered or forming visual images of items) would be beneficial. One suitable training tool would be Memory Booster, mentioned above (see Section 4.3). Students whose recall of verbal information is weak should also be encouraged to use these strategies with other stimuli, for example to practice using rehearsal when remembering a telephone number or a shopping list.

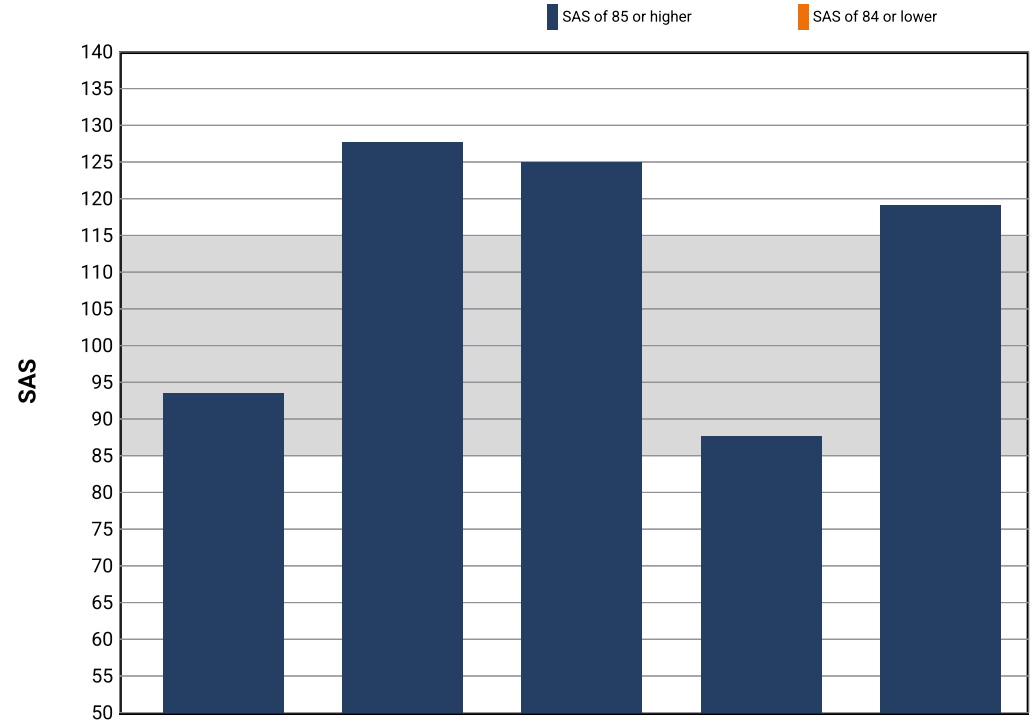
It should also be noticed that Emma's working memory processing speed was close to the lower boundary of the average range, suggesting that she works at a slower rate than most other students. This could impact on her studies as she gets older, especially in examinations and timed assessments. It would be useful for Emma to understand this herself and to be given advice regarding working more quickly.

Recall Individual Student Report

GL Education A Renaissance Company Recall

Name: Emma Watson	**Date of Birth:** 01/11/2013
School: GL School	**Group:**
Date of Test: 03/10/2025	**Age at Test:** 11:11

Results Profile



	Tests				Working Memory
	Word Recall	Pattern Recall	Counting Recall		
	Phonological Loop	Visuo-Spatial Sketch Pad	Executive Function	Processing Speed	Composite
SAS	93	128	125	87	119
Confidence Band	79-107	116-140	113-137	76-98	107-131
Percentile Score	33	97	95	20	90
Raw Score	12	46	27	0.78	46
Memory Span	3	9	6		
Age Equivalent	09:06-10:05	16:00-16:05	16:05	06:08-06:11	14:04-14:07

Assessor's comments

Figure 10. Recall results Emma (age 11:11).

5.3 David (12 years 11 months)

David's performance on Recall is shown in Figure 11.

It is clear that David has poor working memory. An examination of the Recall scores reveals that his performance was poor (below standard score 85) on two of the working memory subtests. This profile of scores is typical of some children with special educational needs, who perform more poorly than age matched controls on measures of the visuo-spatial sketchpad and central executive components of working memory. Therefore, it is likely that David should be recognised as needing extra support for learning, and he is likely to make slow progress with acquiring knowledge and skills in areas such as literacy and mathematics and is at risk of poor educational attainment.

David's results lead to several recommendations. Firstly, it is important that David's teachers recognise that he has a poor working memory, and that his difficulties are not a result of other problems such as inattentiveness or lack of interest in learning. Teachers can then try to reduce the working memory demands of common classroom activities. This involves being mindful that heavy loads are caused by lengthy sentences, unfamiliar content, and demanding mental processing activities. Therefore, where possible teachers should simplify sentences, and use familiar and common words. Teachers should also ensure that David remembers what he is supposed to be doing in any given task, by repeating important information, and asking David to repeat it. David may also benefit from some training related to using visual aids, so that they can be effectively used to reduce the amount of information that needs to be remembered during on-going processing tasks. David should also be encouraged to develop strategies for dealing with poor working memory, including note taking. It is also important to tell David that it is OK to ask for help when it is needed. At nearly 13 years of age, it is also important to realise that David has been coping with a poor working memory for some time. Frequent failures on learning tasks because of a poor working memory may have been detrimental to David's confidence and self-belief. David may therefore benefit from increased support, praise and encouragement during learning activities. Finally, teachers may also want to suggest some kind of working memory training (see Section 4.3). Although the benefits may be somewhat limited, this should at least improve David's confidence and beliefs in his ability to deal with complex processing and storage tasks.

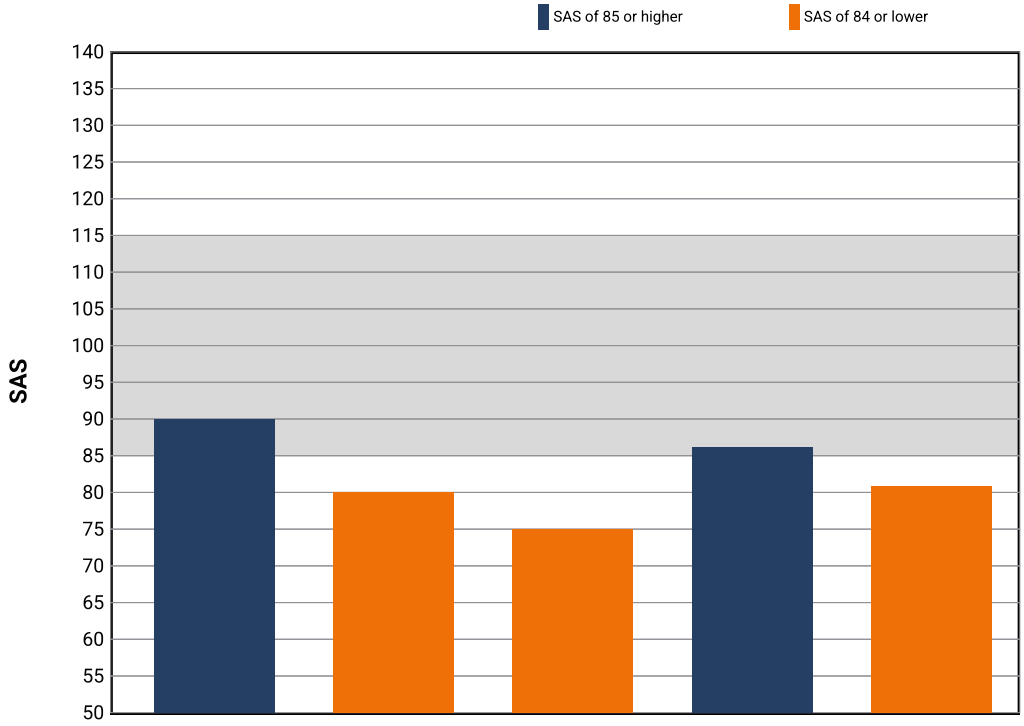
Like the previous case (Emma), David's working memory processing speed was close to the lower boundary of the average range, suggesting that he works at a slower rate than most other students. This could impact on his studies as he gets older, especially in examinations and timed assessments. It would be useful for David to understand this himself and to be given advice regarding working more quickly.

Recall Individual Student Report

GL Education A Renaissance Company Recall

Name: David Fisher	**Date of Birth:** 09/10/2012
School: GL School	**Group:**
Date of Test: 03/10/2025	**Age at Test:** 12:11

Results Profile



	Tests				Working Memory
	Word Recall	Pattern Recall	Counting Recall		
	Phonological Loop	Visuo-Spatial Sketch Pad	Executive Function	Processing Speed	Composite
SAS	90	80	75	86	81
Confidence Band	76-104	<70-92	<70-87	75-97	70-93
Percentile Score	26	9	5	17	10
Raw Score	12	17	2	0.77	21
Memory Span	3	4	2		
Age Equivalent	09:06-10:05	07:00-07:02	05:06-05:11	06:08-06:11	08:03-08:05

Assessor's comments

Figure 11. Recall results for David (age 12:11).

5.4 Rohan (13 years 0 months)

Rohan's Recall scores are shown in Figure 12.

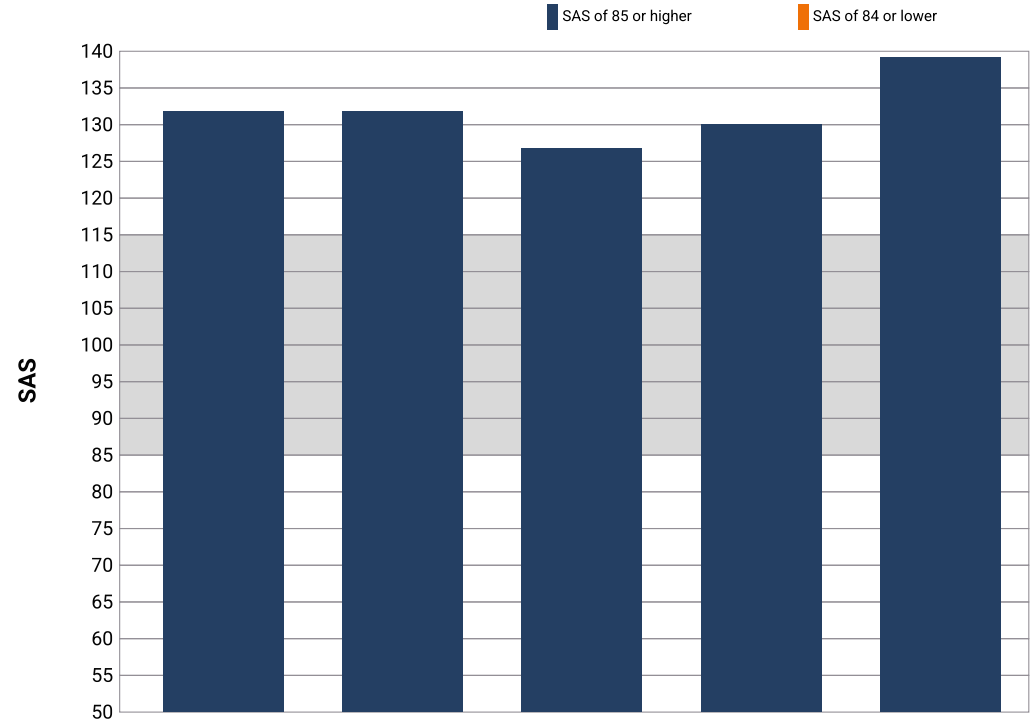
It is clear that Rohan's working memory skills are above average. His performance on each of the three core tests was higher than average for children of this age. His working memory processing speed was also well above average. This suggests that Rohan is good at remembering information in both verbal and visuo-spatial domains and does not have problems with the simultaneous processing and storage of information. Rohan is therefore well equipped for performing well in learning activities and is likely to achieve good levels of educational attainment. If any problems do arise for Rohan, his profile suggests that these will not result from his cognitive skills, but rather, other factors that are important in education, such as motivation and engagement. This is not to say that Rohan would not benefit from interventions such as working memory training or advice on improving study skills. For example, evidence suggests that children with a good working memory still benefit from memory strategy training, although not to the same extent as children with a poor working memory. However, intervention is not seen as critical for children with such a good working memory.

Recall Individual Student Report

GL Education A Renaissance Company | Recall

Name: Rohan Patel	**Date of Birth:** 19/09/2012
School: GL School	**Group:**
Date of Test: 03/10/2025	**Age at Test:** 13:0

Results Profile



	Tests				Working Memory
	Word Recall	Pattern Recall	Counting Recall		
	Phonological Loop	Visuo-Spatial Sketch Pad	Executive Function	Processing Speed	Composite
SAS	132	132	127	130	139
Confidence Band	118->140	120->140	115-139	119->140	127->140
Percentile Score	98	98	96	97	99
Raw Score	25	52	28	0.47	59
Memory Span	6	10	6		
Age Equivalent	16:05	16:06-16:11	16:05	12:06-13:05	>16:11

Assessor's comments

Page 1 of 1

Figure 12. Recall results for Rohan (age 13:0).

5.5 Boris (15 years 10 months) [Examination access assessment]

Boris is in Year 10 and will be sitting GCSE examinations at the end of the year. His literacy skills are average, and he shows good conceptual understanding of material, but he works at a very slow rate so that he rarely completes written exam papers within the time limit. He has been encouraged to increase his rate of working but this turned out to be counterproductive because it dramatically increased the number of errors in his work. Teachers have recognised that when it comes to assessing Boris's skills and knowledge his slow speed of processing disadvantages him substantially and have consequently agreed a school policy to allow him 25% extra time in class written tests and internal exams. The time has now come for him to be assessed for possible access arrangements in forthcoming GCSE exams.

The JCQ regulations that govern procedures for granting access arrangements make provision for students who have slow speed of working, specifically that students with significantly below average performance. The SENCo, who is also the school's qualified assessor for exam access arrangements, assessed Boris using Recall, together with other tests to measure his reading, writing and spelling skills, according to JCQ requirements. His Recall results, shown in Figure 13, indicate that he has poor working memory and slow speed of processing. All his scores, except on Pattern Recall are below standard score 85 and therefore he is eligible for 25% extra time in GCSE examinations.

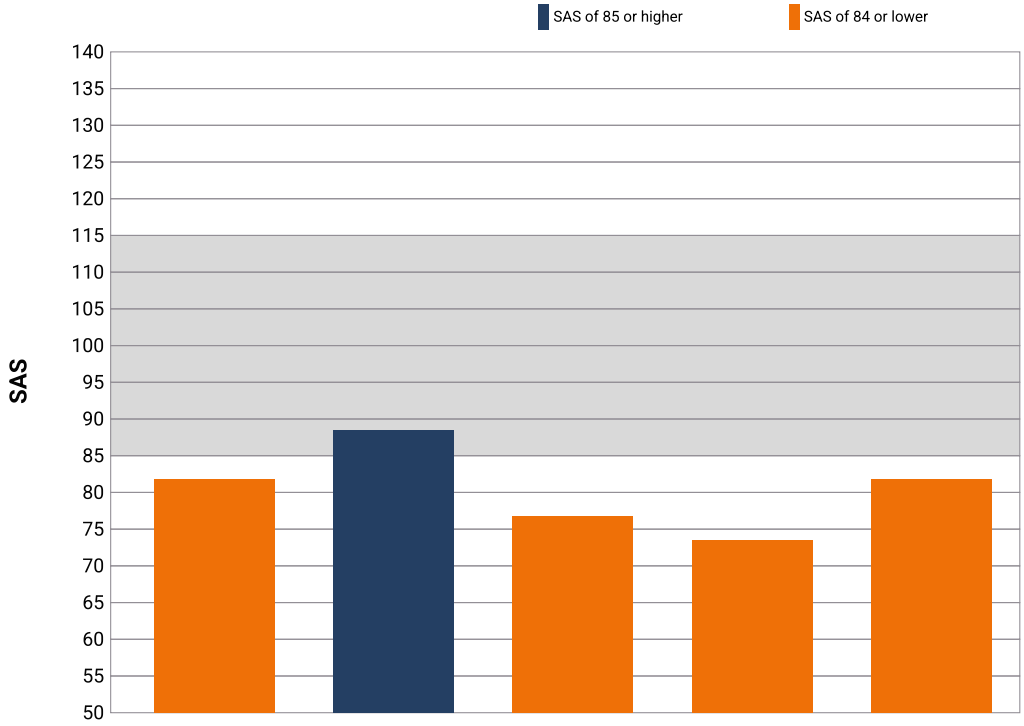
Boris's results were entered on to JCQ Form 8, along with the results of the literacy tests and information about his history of need and the various provisions made for him by the school, as required by JCQ.

Recall Individual Student Report

GL Education A Renaissance Company Recall

Name: Boris Tompkins	**Date of Birth:** 24/11/2009
School: GL School	**Group:**
Date of Test: 03/10/2025	**Age at Test:** 15:10

Results Profile



	Tests				Working Memory
	Word Recall	Pattern Recall	Counting Recall		
	Phonological Loop	Visuo-Spatial Sketch Pad	Executive Function	Processing Speed	Composite
SAS	82	89	77	74	81
Confidence Band	70-96	77-101	<70-89	<70-85	70-93
Percentile Score	12	24	6	4	10
Raw Score	12	33	6	0.82	27
Memory Span	3	7	2		
Age Equivalent	09:06-10:05	11:08-11:11	07:06-07:11	06:00-06:03	09:09-09:11

Assessor's comments

Figure 13. Recall results for Boris (age 15:10).

6 References

Alloway, T.P. (2012). Automated Working Memory Assessment (2nd Edition). UK: Pearson Assessment.

Alloway, T.P., & Gathercole, S. E. (2006). How does working memory work in the classroom? Educational Research and Review, 1, 134- 139.

Alloway, T. P. & Gathercole, S.E. (2006b). *Working memory and neurodevelopmental disorders*. New York, NY: Psychology Press.

Alloway, T.P., Gathercole, S.E., & Kirkwood, H. (2008). Working Memory Rating Scale. UK: Pearson Assessment.

Alloway, T.P., Gathercole, S.E., Kirkwood, H.J., & Elliott, J. (2009). The Working memory rating scale: A classroom-based behavioural assessment of working memory. Learning and Individual Differences, 19, 242-245.

Baddeley, A.D. (2000). The episodic buffer: a new component of working memory? Trends in Cognitive Sciences, 11(4), 417- 423.

Baddeley, A.D., & Hitch, G. (1974). Working Memory. In Bower, G.H. (Ed.), The Psychology of Learning and Motivation. New York: Academic Press.

Case, R., Kurland, D.M., & Goldberg, J. (1982). Operational efficiency and the growth of short- term memory span. Journal of Experimental Child Psychology, 33, 386- 404.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational psychology review*, *26*(2), 197-223.

Della Sala, S., Gray, C., Baddeley, A.D., & Wilson, L. (1997). The Visual Patterns Test: A test of short-term visual recall. Thames Valley Test Company.

Elliott, C.D. (1983). The British Abilities Scales. Windsor: NFER Nelson.

Gathercole, S.E., & Alloway, T.P. (2008). Working memory and learning: A practical guide for teachers. London: Sage.

Gathercole, S.E., Brown, L., & Pickering, S.J. (2003). Working memory assessments at school entry as longitudinal predictors of National Curriculum attainment levels. Educational and Child Psychology, 20, 109- 122.

Gathercole, S.E., & Pickering, S.J. (2001). Working memory deficits in children with special educational needs. British Journal of Special Education, 28, 89-97.

Gathercole, S.E., Pickering, S.J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. Applied Cognitive Psychology, 18, 1- 16.

Horne, J.K. (2007) Gender differences in computerised and conventional educational tests. Journal of Computer Assisted Learning, 23, 47–55.

Juffs, A. and Harrington, M. (2011) Aspects of working memory in L2 learning. Language Teaching, 44, 137-166.

Kane, M.J., Brown, L.G., McVay, J.C., Silvia, P.J., Myin-Germeys, I., & Kwapil, T.R. (2007). For whom the mind wanders and when: An experience sampling study of working memory and executive functions in everyday life. Psychological Science, 18, 614-621.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental psychology*, *49*(2), 270.Pickering, S., & Gathercole, S. (2001). Working Memory Test Battery for Children (WMTB- C). London: The Psychological Corporation.

Rapport, M. D., Orban, S. A., Kofler, M. J., & Friedman, L. M. (2013). Do programs designed to train working memory, other executive functions, and attention benefit children with ADHD? A meta-analytic review of cognitive, academic, and behavioral outcomes. *Clinical psychology review*, *33*(8), 1237-1252.

Redick, T. S., Shipstead, Z., Wiemers, E. A., Melby-Lervåg, M., & Hulme, C. (2015). What's working in working memory training? An educational perspective. *Educational Psychology Review*, *27*(4), 617-633.

Sana, F., & Fenesi, B. (2025). Working Memory and Instructional Fit: Reintroducing Aptitude–Treatment Interaction in Education Research. *Behavioral Sciences*, *15*(6), 765.

Shipstead, Z., Redick, T.S., & Engle, R. W. (2012). Is working memory training effective? Psychological Bulletin, 138, 628- 654.

Singleton, C.H. (2001) Computer-based assessment in education. Educational and Child Psychology, 18, 58-74.

St Clair-Thompson, H.L. (2011). Executive functions and working memory behaviours in children with a poor working memory Learning and Individual Differences, 21, 409-414.

St Clair- Thompson, H.L. & Gathercole, S.E. (2006). Executive Functions and Achievements in School: Shifting, Updating, Inhibition, and Working Memory. Quarterly Journal of Experimental Psychology A, 59, 745-759.

St Clair-Thompson, H.L., Stevens, R., Hunt, A., & Bolder, E. (2010). Improving children's working memory and classroom performance. Educational Psychology, 30, 203-220.

Wechsler, D. (1986). Wechsler Intelligence Scale for Children- Revised. New York: Psychological Corporation.