A Guide to Standardised Assessments

Risingstars-uk.com/assessment
An introduction for school leaders and teachers

What are standardised assessments and why might a school use them?
How are standardised assessments developed and how do they work?
What information can standardised assessments give you?
How can data from standardised assessments support your school’s teaching and learning?

Look out for this icon for specific information on our termly standardised test suites PiRA, PUMA, GAPS and NTS Assessments
What are standardised assessments and how are they used?

What are standardised assessments?

A standardised assessment is any form of test that requires all test takers to answer the same questions in the same way and is then scored in a ‘standard’ or consistent manner. This makes it possible to compare the relative performance of individual pupils or groups of pupils. The term is primarily associated with large-scale tests administered to large populations of pupils.

Standardised assessments follow a rigorous development process involving trialling the tests with a nationally representative sample of children, followed by statistical analysis of the results. This allows schools to then benchmark their pupils’ performance against the nationally representative group of children used in the standardisation trial.

Why use standardised assessments?

Standardised assessments are very useful when schools want to know how their pupils are performing relative to pupils in other schools. Nationally standardised tests provide information on how pupils are performing over time and in comparison to pupils nationally. They help teachers understand national expectations and assess their pupils’ performance in the broader national context.

In addition, standardised assessments enable school leaders and governors to benchmark their school’s performance against other schools within a group or cluster and nationally.

How do standardised assessments work?

In order to produce reliable results, standardised tests must be administered and marked according to the test guidance. In particular, the tests must be taken at a fixed point, for example at the beginning or end of a specified school year. Failure to follow these arrangements can affect the reliability of the results.

In the case of PiRA, PUMA, GAPS and NTS Assessments, they are designed for use on a termly basis, in the second half of each term.

What are the limitations of standardised assessments?

As with all tests, standardised assessments don’t measure everything a pupil has been taught in a particular subject area. In addition, they are a snapshot at a point in time and may not be a true reflection of what a pupil knows. For example, they may be feeling unwell on the day or may not perform in line with expectations under the pressure of test conditions. For these reasons, standardised tests are best used alongside regular in-class formative assessment, which provides teachers and pupils with valuable, real-time information about what pupils know and can do and what they need to do to improve.
What are the benefits of standardised assessments?

Standardised assessments are perceived as being fairer than non-standardised assessments because all pupils take the same test and the mark scheme is clearly defined. This consistency also allows a more reliable comparison of outcomes across all test takers.

The test results also give an indication of how a pupil (or group) is performing in relation to pupils in other schools. The use of standardised tests over time enables progress to be tracked in an independent way.

Standardised tests can be used before and after an intervention to help measure the impact that it’s had.

The reading, mathematics, grammar, punctuation and spelling ages generated by our standardised assessments are ideal for monitoring the impact of interventions.

How are standardised assessments developed?

Standardised assessments are developed in a rigorous way to ensure that they are both valid (the test measures what it has been designed to assess) and reliable (the test results are consistent, dependable and replicable).

This process typically takes at least 18 months and involves:

• developing a robust test framework that defines the content or skills to be assessed
• writing and reviewing the test questions
• trialling the questions and tests with a representative sample of pupils (at least 1,000 per year group). Trials take place at the time the tests are designed to be used

For PIRA, PUMA, GAPS and NTS Assessments, trialling took place in the second half of the Autumn, Spring and Summer terms.

• statistically analysing these trial results to produce standardised scores and other information so that the results can be interpreted and used by teachers in a meaningful way
• producing the final test(s) on the basis of triallist feedback and the outcomes of the trial(s).
What information is produced from standardised tests and what does it tell you?

The scores are all independent of each other. They provide different information about a pupil’s performance and you should use the scores that best meet your needs and those of your school.

The following scores are the main numerical information provided from standardised tests. Their meaning is the same regardless of the test provider.

**Raw score**

This is the total number of marks that a pupil has scored in a particular test.

It is used as the starting point for all of the other scores.

A table of average raw scores for each PiRA, PUMA, GAPS and NTS Assessments test by term and gender is provided in the test guidance so that you can compare how well each child is doing in relation to the average in the standardisation sample. This can give you a quick overview of where each pupil is and also enables you to compare overall performance of boys and girls.

**Standardised score**

This is derived from the pupil’s raw score and placed on a scale that makes a comparison with the nationally representative sample of pupils.

Standardised scores are useful for comparing the performance of one cohort with another, provided they have taken the same test and it has been administered correctly.

Test scores are standardised to an average score of 100, making it easy to see whether a pupil is above or below the mean, relative to the national standardisation sample.

A standardised score of 100 is the average test score; it is not the same as a National Test scaled score of 100, which is set by the DfE in advance of testing as the expected standard a pupil should achieve at the end of KS1 and KS2. On National Tests, the raw score required to get a scaled score of 100 may change year-on-year. On a standardised test, the raw score required to get a standardised score of 100 does not change because it is based on the results of the trial.
Percentile score

Percentiles can give you a better feel for the significance of a pupil’s performance because they show the percentage of pupils obtaining any score or lower.

A percentile of 50 is average and a score that is in the 68th percentile means that 68% of the group scored below that pupil’s standardised score. The child is therefore in the top-third for his/her cohort.

Age-standardised score

This is also based on the pupil’s raw score, adjusted for age – in months – and placed on a scale that makes a comparison with those involved in the standardisation trial.

This allows you to compare the performance of pupils against others of the same age. The benefit is that younger pupils in a class are not disadvantaged by better performance of older pupils in their year group. In addition, age-standardised scores enable you to compare performance in any age-standardised test. Note, however, that age is less likely to have an impact on performance in standardised tests that are designed to assess curriculum knowledge and understanding (like PiRA, PUMA, GAPS and NTS Assessments) because all pupils in the year group are being taught the same content. In curriculum tests, age generally has less of an impact on performance in Key Stage 2 than Key Stage 1.

Reading age/Mathematics age/ GPS age

Get a quick measure of attainment against the age at which a pupil’s performance is typical.

The reading, mathematics or GPS (grammar, punctuation and spelling) age can be used as a quick indication of a child’s reading, maths or GPS ability. It is useful for monitoring the impact of interventions, as well as being a good measure for pupils entering a school from other schools. Please note that reading, mathematics or GPS ages are not the same as reading, mathematics or GPS attainment.

PiRA, PUMA and GAPS also include:

Hodder Scale

The Hodder Scale is a decimal scale ranging from 0 to 6.0, which enables progress to be monitored term-by-term and year-by-year. In addition, it can be used as a basis for estimations about future progress.

The Hodder Scale acts as a common ‘spine’ on which all of the PiRA, PUMA and GAPS tests across the primary phase are plotted. It was derived from the raw scores of over 1,000 pupils in each year group in the standardisation trial. The main benefit of the Hodder Scale is that it provides a whole-school way of monitoring, reporting and predicting progress in the absence of National Curriculum levels.

NTS Assessments also include:

NTS Scale

The NTS Scale has its origins in the Hodder Scale but is specific to NTS Assessments. Like the Hodder Scale, it is an independent decimal scale that can be used to track progress and estimate future performance. It is particularly useful for children who are working outside of their age range but who continue to make small increments of progress for their year.
Case study 1

In Year 2, Tarren was tested using PiRA and he started the year well, scoring 18 (a standardised score of 103) in the Autumn test, the equivalent of 2.2 on the Hodder Scale. However, in the Spring term he fell back to a standardised score of 95, which meant he stayed at 2.2 on the Hodder Scale. Although the drop was not significant, it corroborated the class teacher’s report that Tarren seemed to be losing confidence.

At the end of the Spring term, his teacher looked back at his previous PiRA tests and noticed that Tarren was gaining all his marks from the easy questions and seemed to hit a barrier when it came to the harder ones. Tarren’s teacher decided that for the Summer term he would move Tarren to a higher literacy group, in order to give him more experience of answering harder questions in the supportive context of a guided reading group. The Summer term results were much more encouraging (scoring 14 marks, 2.5 on the Hodder Scale and a standardised score of 96). Tarren gained confidence as a reader and was showing progress, albeit his marks showed that he was still below the average standardised score of 100 and he still had some catching up to do.

Case study 2

In the NTS Assessments Year 5 Summer test, Sofia had a standardised score of 110 and an NTS Scale value of 4.5, giving an estimated NTS Scale value of 4.8. However, when she sat the NTS Assessments Year 6 Autumn test, she only achieved a standardised score of 101 and an NTS Scale value of 4.5 again, indicating that her progress had stalled. Sofia’s teacher was concerned because her progress had previously been steady. She reviewed Sofia’s scores and there was no obvious weak strand so she looked in further detail at Sofia’s test papers.

The teacher discovered that the more complex arithmetic items were causing Sofia problems under timed conditions. She decided to have a chat with Sofia, which revealed that she had recently started getting worried about the SATs and this was causing her to panic in tests.

The teacher then checked the complex arithmetic marks, specifically long multiplication and long division, for the rest of the class. This proved to be a difficult strand for many learners, so the teacher adjusted her lesson planning to focus on this area. She is delighted that she can analyse her pupils’ progress at such a granular level and is able to use the data to benefit all her pupils, not just the outliers at each end of the ability spectrum.
Q ‘I want to track my pupils’ progress from Autumn through Spring and Summer in Year 3.’

A Use the Hodder Scale or NTS Scale to monitor progress term-by-term and year-by-year.

Q ‘How well are our Summer-born Year 1s doing so far this year?’

A Use the age-standardised score to see how pupils are performing against other pupils of the same age.

Q ‘I want to determine the effectiveness of a mathematics resource used with my intervention group.’

A Use their mathematics ages before and after the use of the resource to monitor impact.

Q ‘I would like to see how well we are performing as a school in reading.’

A Use the standardised scores to compare performance on a national scale.

The PiRA, PUMA and GAPS termly standardised assessments enable schools to accurately measure and predict pupil progress in reading, mathematics, grammar, punctuation and spelling, and benchmark their performance against national averages. NTS Assessments offer the same for reading and mathematics but with National Test-style papers.

Whichever suite you choose, with three tests for each year group, the results provide reliable diagnostic information to support and guide effective teaching and learning, and offer evidence of progress for parents, governors and Ofsted.