



NFER

Classroom



Assessment 101

A trainee and early career teacher's handbook
to primary assessment in England

For our full range of support on assessment in schools, visit
the NFER Assessment Hub at www.nfer.ac.uk/assessment-hub

3

What is assessment?

3

Summative assessment

4

Formative assessment

5

Principles of assessment

6

Statutory assessments in key stage 1 and key stage 2

6

Key stage 1

7

Key stage 2

8

Optional tests in key stage 1 and key stage 2

8

Standardised tests

9

Ofsted and assessment

10

Glossary

Contents

What is assessment?

Many people assume that 'assessment' means taking a test, but assessment is broader than that. There are two main types: summative and formative. It is likely that both happen at some level in most classrooms, and both can facilitate good teaching. The key is to understand what each method can contribute and to build your practice to maximise the effectiveness of each.

Summative assessment

Summative assessment is what we usually think of when we consider assessment. It 'sums up' what a pupil has achieved at the end of a period of time or block of learning, and is usually linked to learning aims and any relevant national standards. There may be an assessment at the end of a topic, at the end of a term or half-term, at the end of a year or, as in the case of the [national curriculum tests](#), at the end of a key stage. A summative assessment may be completed independently by a pupil, or may be conducted by another person, such as the teacher. It may be a written test, an observation, a conversation or an active task. It may be recorded in writing, through photographs or other visual media, or through an audio recording. Some summative assessment is completed using software on a computer or laptop. Whichever medium is used, the assessment will be marked either automatically or by a person, and a score, grade or other indicator will be derived for the pupil.

Scores derived from summative assessment are used in many ways. Some of these include:

- to measure the pupil's skills, knowledge or understanding of a particular topic
- to monitor the progress a pupil has made over time
- to inform parents of a pupil's progress
- to inform other stakeholders (governors, school leadership).

Summative assessment is often a gateway to the next level of learning, or to achieving a qualification. Sometimes, a summative assessment is statutory for pupils, but the results are only used to reflect on the school, not the pupil. It's important to remember that summative assessment has a specific purpose, and the outcomes aren't intended to be indicative of everything a pupil knows or can do. Summative assessment outcomes should therefore inform rather than determine teachers' own assessment decisions.

Formative assessment

Formative assessment, or assessment for learning, is the method of integrating less formal, iterative assessment into the learning process, to allow both pupils and teachers to shape the learning process depending on outcomes. It is widely agreed to be a crucial element of successful teaching, and should take place on a day-to-day basis in the classroom.

It often begins with [diagnostic assessment](#), which is an early assessment where the teacher gathers a picture of what the pupil already knows and what gaps may exist in their skills, knowledge or understanding. If the teacher and pupil understand what has been achieved to date, it is easier to plan the next steps. As the learning journey unfolds,

further formative assessments indicate whether teaching plans need to be amended to reinforce or extend learning. Formative assessments may be questions, conversations, observations, tasks, quizzes or more formal assessments, and may be conducted teacher-pupil, teacher-class, or pupil-pupil. Often formative assessment may not be formally recorded at all, except perhaps in the lesson plans drawn up to address the next steps indicated.

The assessment for learning approach recognises that every classroom activity offers assessment opportunities for teachers to evaluate learning, provide feedback and inform their teaching.

Some key features of assessment for learning are:



Explaining learning objectives and [success criteria](#) to pupils



Creating informal opportunities to assess pupils' knowledge and understanding



Creating an environment in which pupils can learn from each other



Using questions and dialogue that promote deep learning



Giving time for learning to be absorbed



Giving constructive feedback



Encouraging pupils to reflect upon their learning and to monitor their own progress



Using assessment to inform future actions

Principles of assessment

VALIDITY

This is the overarching assessment principle. It concerns how assessment outcomes are interpreted and used. An assessment may be reliable and fair but it may still be inappropriate to use it in a particular context. For example, a teacher may wish to identify pupils who need support in structuring their ideas in writing. To give the whole class a spelling test and use the results to identify the pupils who would benefit from additional support in organising their writing is not a valid use of the assessment. The spelling test may be a well-developed assessment with good measurement properties but it is not valid for this purpose.

For more on this, check out our short-read on the Assessment Hub: '[How to ensure a test is valid](#)'.



RELIABILITY

The principle of reliability in assessment refers to the importance that assessment outcomes are consistent. A consistent approach to administration ensures that the results are not dependent on the specific assessor; similarly mark schemes or performance criteria are designed to be interpreted in the same way by different people. In measurement, the internal consistency of an assessment refers to how effectively separate items or parts of an assessment are measuring the same construct.



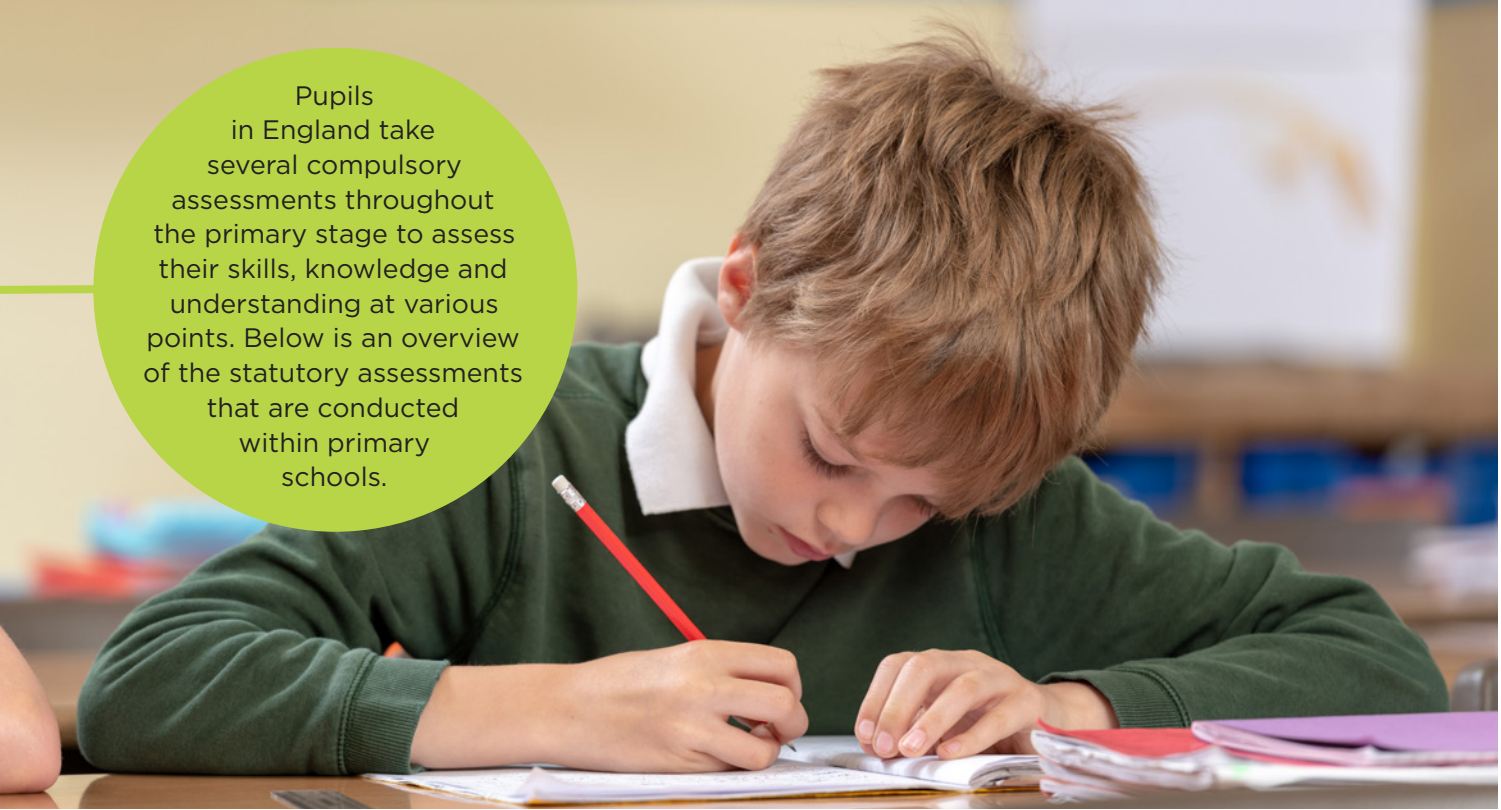
FAIRNESS

There are two key aspects of this principle. One relates to the underlying construct being assessed and the impact of other, unrelated, attributes. The other relates to the impact of prior experience.

It is important that pupils can demonstrate their abilities in the construct (e.g. the skill or knowledge) being assessed. Access arrangements or accommodations may be available in an attempt to 'level the playing field'. The aim of these is to reduce the impact of factors which may affect performance but are not relevant to what is being assessed. An example is when print is enlarged to assist those with a visual impairment. In some instances, we may be able to introduce an access arrangement in one subject but not another. For example, it is possible to have a reader in a mathematics test but this would be inappropriate in a reading test.

The other, equally important, element of fairness is to ensure that the assessment content does not unfairly advantage or disadvantage some groups in the population being assessed. An example here would be asking candidates to discuss the introduction of a variation of the offside rule in a speaking and listening assessment. This would potentially disadvantage those unfamiliar with the laws of football.

Statutory assessments in KS1 and KS2



Pupils in England take several compulsory assessments throughout the primary stage to assess their skills, knowledge and understanding at various points. Below is an overview of the statutory assessments that are conducted within primary schools.

Key stage 1

National curriculum assessments

National curriculum assessments are the statutory assessments carried out in primary schools in England, sometimes colloquially referred to as 'SATs'. These assessments are used to determine whether pupils have mastered the skills, knowledge and understanding set out in the national curriculum programmes of study by the end of each key stage. Pupils are given a scaled score, with 100 representing the expected standard.

At the end of key stage 1 (year 2), pupils are required to take national curriculum tests in mathematics and English reading, which are delivered and marked by teachers in May each year. Teachers are required to use these test results in conjunction with evidence from their own ongoing assessment of pupils throughout the year to provide a summary of attainment for each pupil. Pupil achievements are benchmarked against a framework of criteria provided by the Standards and Testing Agency. Although there are no tests for English writing and science in key stage 1, teachers are also required to assess and report pupils' knowledge and understanding of these subjects through teacher judgements at the end of the key stage.

Teacher assessment outcomes are collected and the national data is reported annually. Teacher assessment outcomes must also be reported to parents. School-level results are made available to Ofsted.

There is also an optional KS1 spelling, punctuation and grammar test which schools can choose whether or not to administer. Following the anticipated introduction of the [Reception Baseline Assessment](#) in 2021, it is likely that the KS1 national curriculum tests in English reading and maths will also be made non-statutory, meaning that schools will be able to choose whether or not they wish to administer them.

Phonics screening

The phonics screening check (PSC) is a compulsory assessment taken by all children in Year 1 attending a state school in England. The purpose of the PSC is to confirm that all children have learned phonic decoding to an age-appropriate standard and identify those that require extra support from their school to improve. Individual children's results are made available to parents so that they are kept informed about their child's progress in developing word-reading skills. School-level results are recorded and made available to Ofsted for use in inspections but are not published in performance tables.

Key stage 2

National curriculum assessments

At the end of KS2 (year 6), all pupils (except those for whom the curriculum is disapplied) take national curriculum tests in mathematics, English reading, and English grammar, punctuation and spelling. The tests are marked by trained markers – not teachers – and pupils are given a scaled score for each subject to show whether they have reached the expected standard.

For English writing, pupil attainment is assessed and reported by the teacher at the end of KS2. Teachers need to base their judgement on a broad range of evidence, which will come from day-to-day work in the classroom. Teacher assessment frameworks are made available each year by the Department for Education (DfE).

Science is also reported by teacher assessment alone. However, once every two years, a nationally representative sample of schools participate in the science sampling tests, which provide a national picture of science attainment.

Individual pupil results from national curriculum assessments are reported to parents in the form of scaled scores. National results are published before the end of the summer term. School-level results are made available to Ofsted and also appear in performance tables.

Multiplication tables check

The multiplication tables check (MTC) is an online assessment for pupils to complete at the end of year 4 either on a computer or a tablet to determine whether pupils can fluently recall their multiplication tables. The test includes 25 questions and will take five minutes. A total score out of 25 will be reported to each school for all of their pupils who took the check.

There is no expected standard threshold for the MTC, beyond the number and percentage of pupils who achieve full marks, as fluency is the focus of the assessment. Individual pupil and school-level results are made available to schools to allow them to provide additional support to pupils who require it. School-level results are available to Ofsted via the Analyse School Performance (ASP) data system but they will not be published in performance tables.

RESOURCES

More information about KS1 and KS2 assessment, including sample test materials and teacher assessment frameworks, can be found on the [DfE website](#).



National curriculum assessments are used to determine whether pupils have mastered the skills, knowledge and understanding set out in the national curriculum programmes of study by the end of each key stage.

Optional tests in KS1 and KS2

Many schools choose to administer additional tests at certain points throughout the school year to provide an independent measure of attainment and progress. Whereas statutory national assessments are used primarily as an accountability tool, a key benefit of optional assessments is that they can be used diagnostically to identify pupil strengths and weaknesses to inform ongoing teaching, i.e. as formative assessment.

Standardised tests

Many people will remember test scores from their school days such as '7 out of 10'. These are known as raw scores, are readily understandable and useful in indicating what proportion of the total marks a pupil has gained in a test. However, these scores are less useful in enabling teachers to compare pupils' performance meaningfully between one test and another, and to monitor progress over a period of time. For this reason, when opting to assess pupils in addition to statutory testing, many schools choose standardised tests.

A standardised test is scored in a "standard" or consistent manner which makes it possible to compare the relative performance of pupils. The test is standardised: trialled with a nationally representative sample of children before being released, so that performance can be compared and amalgamated to provide national benchmarks of performance. Usually, standardised tests are constructed so that the

average, nationally standardised score is 100, irrespective of the difficulty of the test. This means teachers can readily identify whether a pupil is above or below the national average.

Standardised tests are a popular choice among schools as, when robustly developed and correctly administered to pupils, they can produce a reliable measure of pupil attainment and progress. Not only do they allow schools to benchmark pupil performance nationally, they enable meaningful comparisons between pupils and groups of pupils and provide a measure of attainment that can be monitored over time to indicate progress.

NB: Guidance from the DfE's Workload Advisory Group suggests that there is no need for more than one attainment data collection point per term, which should be used to inform clear actions.

Ofsted and assessment

In May 2019, after a large-scale consultation with schools and stakeholders, Ofsted published its new Education Inspection Framework, a document outlining the way in which it inspects schools.



Among the changes made in the new framework, which came into effect in September 2019, was that Ofsted will not consider school's internal assessment data as evidence of 'impact' (i.e. pupil achievement) during inspections. Instead, inspectors will rely solely on national performance data alongside first-hand evidence gathered during inspections as evidence of the quality of education in terms of its impact on pupils.

However, this is not to say Ofsted does not want to see evidence of the use of internal assessment at all. On the contrary, the new framework states that schools' use of assessment will be considered as evidence of the quality of education in terms of 'implementation'. More specifically, inspectors will consider the extent to which 'Teachers and leaders use assessment well, for example to help learners embed and use knowledge fluently or to check understanding and help inform teaching.'

The framework goes on to state that the quality of education provided at a school will be judged on whether teachers 'check learners' understanding systematically, identify misconceptions accurately and provide clear, direct feedback,' and 'in doing so, [that] they respond and adapt their teaching as necessary, without unnecessarily elaborate or differentiated approaches.'

Internal assessment, including standardised testing and assessment for learning techniques, can be extremely valuable in helping learners embed and use their knowledge, as well as providing a diagnostic tool to enable schools to monitor learning in order to inform ongoing teaching. Find out more about diagnostic assessment on the [NFER Assessment Hub](#).

To summarise, Ofsted will not be judging the results from internal assessment as proof of pupil attainment, but will be interested in seeing how schools use internal assessment to further pupil learning. Schools may wish to use reports generated by online analysis tools to effectively communicate their use of assessment with inspectors, however these will not be requested during inspections.



Ofsted will be interested in seeing how schools use internal assessment to further pupil learning.

Glossary

Sometimes a good old glossary of terms can be the most valuable resource when getting to grips with a new area, so here you can find definitions for some of the most common terms used in relation to assessment.



ACCOUNTABILITY

A government's mechanism for holding educational institutions to account for the delivery of high quality education.

AGE-RELATED EXPECTATIONS

Standards that define what is expected of a pupil by a specified age or year group. Read more about [age-related expectations](#) on the Assessment Hub.

AGE-STANDARDISED SCORES

Standardised scores that account for the age of a pupil.

ANALYSE SCHOOL PERFORMANCE (ASP) SYSTEM

Replaced RAISEonline in England in 2018 and provides detailed performance analysis to schools.

ASSESSMENT COORDINATOR/ASSESSMENT LEADER

The member of school staff who leads on assessment.

ASSESSMENT FOR LEARNING

Any assessment activity that guides further learning. See Formative assessment.

ASSESSMENT LITERACY

Knowledge and understanding of good assessment practice.

BASELINE ASSESSMENT

Used to establish a point from which future measurements and predictions can be calculated. By comparing a pupil's performance with their baseline assessment, attainment and progress can be monitored.

.....

COMPARATIVE JUDGEMENT

An assessment technique which involves comparing a series of two pieces of work side-by-side to establish a measurement scale. The judge focuses on identifying which of the pieces is of higher quality.

.....

COMPUTER-ADAPTIVE TEST

A computer-based test that presents easier or more difficult tasks based on previous answers (i.e. it adapts to the test-taker's performance). Sometimes referred to as 'CAT tests' or 'personalised tests'.

.....

COMPUTER-BASED ASSESSMENT

A test, task or other assessment activity completed on a computer, sometimes, but not necessarily, delivered online. Also known as E-Assessment.

.....

CONTEXTUAL VALUE ADDED

A value added measure of the effectiveness of a school or the progress made by individual pupils, taking account of pupils' starting points (prior attainment) and school and pupil contextual factors (such as the number of pupils in the school eligible for free school meals or the number of pupils for whom English is not their first language).

.....

CRITERION-REFERENCED ASSESSMENT

An assessment where pupils are assessed against a criterion or set of criteria rather than evaluating them in comparison with the performance of other pupils. The criteria represent a level of expertise or mastery of skills or knowledge (in theory all pupils could meet or fail the criteria.) The driving test is an example of a criterion-referenced test.

.....

CURRICULUM TEST

An assessment of what children have learned, rather than a test of skill. The assessment is based upon content from a pre-defined curriculum or programme of study.

.....

DATA (ASSESSMENT)

Outputs from assessment, such as test results, that can be collected and recorded for reference or analysis.



DIAGNOSTIC ASSESSMENT/ DIAGNOSTIC TEST

An assessment that allows a teacher to determine pupils' individual strengths, weaknesses, knowledge, and skills prior to a new period of learning. For more on the benefits of diagnostic assessment, check out our '[Ask the expert](#)' piece on the NFER Assessment Hub.

.....

EARLY YEARS FOUNDATION STAGE PROFILE

An assessment used to summarise outcomes and progress at the end of the foundation stage of education.

.....

E-ASSESSMENT

A test, task or other assessment activity completed on a computer, sometimes, but not necessarily, delivered online.

.....

EVALUATION

Gathering evidence from a range of sources that inform on performance and using this to support improvements in practice.

.....

FORMATIVE ASSESSMENT

Assessment activities that show teachers where pupils are in their learning and help teachers decide what to teach next (sometimes referred to as assessment for learning).

.....

HIGH-STAKES TESTING

Typically refers to national assessment programmes. The term 'high-stakes' refers to the significance of the outcomes for individual pupils (e.g. secondary selection tests) or for schools (accountability judgements about school performance).

.....

LEAGUE TABLES

See performance tables.

.....

NATIONAL CURRICULUM TESTS (NCTS)

Statutory tests developed by the Standards and Testing Agency (STA) for primary schools in England (often referred to as SATs). Pupils take externally-marked national curriculum tests in reading, grammar, punctuation and spelling, and mathematics at the end of key stage 2 (end of year 6). Pupils take national curriculum tests to support teacher assessment at the end of key stage 1 (end of year 2). One of the purposes of the tests is to judge school performance.

.....

NORM-REFERENCED ASSESSMENT

An assessment where pupils are assessed in comparison with the performance of other pupils (e.g. standardised tests). 'Norm' groups (e.g. the samples with which pupils are compared) are usually ones in which the scores are distributed in a 'normal' or 'bell-shaped' curve.

.....

OFF THE SHELF TESTS

Commercially published tests which can be purchased by schools. Some – but not all – will be standardised tests.

.....

OPTIONAL TESTS

Non-statutory tests used to assess pupils' knowledge, understanding and learning.

.....

PEER-ASSESSMENT

Process of pupils taking responsibility for assessing the work of their peers.

.....

PERCENTILE RANKS

Refer to a pupil's performance in relation to the 'norm' (i.e. other pupils who took the same test). If a pupil's percentile rank is 80, it means that he/she achieved a score equal to or better than 80 percent of the pupils who took the test.

.....

PERFORMANCE TABLES

Also known as league tables. These summarise how well a school is doing, using a number of performance measures, including the percentage of pupils achieving the 'expected standard' in reading, writing and mathematics; pupils' average scaled scores in reading and mathematics and pupils' average progress in reading, writing and mathematics at the end of key stage 2.

.....

PHONICS SCREENING CHECK

An assessment designed for use at the end of year 1 to confirm whether pupils have learnt phonic decoding to an appropriate standard. The check consists of 20 real words and 20 pseudo-words that a pupil reads aloud to the teacher.

.....

(INTERIM) PRE-KEY STAGE STANDARDS

For the statutory assessment of pupils who are working below the overall standard of national curriculum tests and are engaged in subject specific learning. See also P scales.

.....

PUPIL PREMIUM

An additional funding given to schools so that they can support disadvantaged pupils and close the attainment gap between them and their peers.

.....

RAW SCORE

The actual mark or score obtained by a pupil in a test. This may be converted to a Standardised score.

.....

RECEPTION BASELINE ASSESSMENT

A baseline assessment of children's attainment (primarily early literacy and numeracy skills) on entry to a reception class. A new assessment is currently in use by early adopter schools and due to become statutory in reception in Autumn 2021. Its purpose will be to act as the starting point for calculating progress when pupils reach the end of key stage 2.

.....

RELIABILITY

A statistical measure of internal consistency in a test. A high reliability value (the maximum is 1) suggests that the questions in a test are related and therefore assessing the same thing. Other measures of reliability exist, e.g. indicating test stability and whether different versions of a test are interchangeable (i.e. the extent to which repeating the test or using different versions of a test would give similar results).

.....

REPRESENTATIVE SAMPLE

A sample group that accurately represents the composition of the population group of interest. Representative samples of test-takers are used in assessment trials so that valid and reliable conclusions can be drawn about the general population of test-takers.

.....

SATS

See national curriculum tests.

.....

SCALED SCORE

Converted from raw scores, scaled scores enable accurate comparisons of performance over time. In national curriculum tests a scaled score of 100 always represents the expected standard on the test. The NCT scaled scores are different from standardised scores (see below) where the 100 represents the average at the time of standardisation.

.....

SELF-ASSESSMENT

Process of pupils taking responsibility for assessing their own work.

.....

STANDARD DEVIATION

A measure that is used to quantify the amount of variation or dispersion of a set of data values. In assessment, it quantifies the average number of points between all test scores and the average score.



STANDARDISED TESTS

Published tests that have been trialled on a nationally representative sample and that can therefore be used for benchmarking individual/group performance against the national average. Schools can choose whether or not to use these tests and how frequently. For example, NFER publishes [termly standardised tests](#) in reading, mathematics, grammar and punctuation and spelling for years 3 to 6 and tests of reading and mathematics for years 1 and 2.

.....

STANDARDISATION

The process of making something conform to a standard. In assessment, tests can be standardised through trials on a large nationally representative sample so a pupil's performance (see Standardised score) can be benchmarked against the national average, and meaningfully compared with other pupils and standardised scores from other tests.

.....

STANDARDISED SCORE

A score that is converted onto a common scale so that the achievement of pupils can be compared directly. Useful for comparing attainment of pupils who took different versions of a test or for monitoring relative progress. Age-standardised scores allow for differences in the age of pupils taking the same test.

.....

STANDARD SETTING

The process of establishing cut scores on an assessment. This helps to create categories such as pass / fail or grades 9-1. A variety of statistical and judgemental methods can be used to set standards.

.....

STATUTORY ASSESSMENTS

Assessments that schools are legally obliged to carry out (e.g. at the end of a key stage).

.....

SUCCESS CRITERIA

The evidence a teacher looks for when deciding whether a pupil has successfully learned something. For example, for a pupil beginning to learn about using full stops, the teacher's success criteria might be that the pupil should use a full stop at the end of all sentences in a piece of writing that is one page long.

.....

SUMMATIVE ASSESSMENT

A type of assessment used at the end of a topic, year or phase of education to show what pupils have learned.

.....

TARGET

Defines what pupils will aim to achieve next. Targets may be short term or longer term. They may be set for individual pupils, for groups, or for a whole class.

.....

TASK

A practical activity used for assessment purposes. It may be observed by a teacher who will take notes on how the pupil completes the task.

.....

TEACHER ASSESSMENT FRAMEWORKS

Guidance from DfE in England to be used to support statutory teacher assessment judgements at the end of key stages 1 and 2.

.....

TEST

A series of questions that a pupil answers on his or her own, without help. Generally answered on paper or on a computer.

.....

VALIDITY

Reflects the extent to which an assessment enables correct inferences to be made about a pupil's knowledge, understanding and skills in the subject being assessed; i.e. the extent to which the test is appropriate for a given population and purpose. Validity judgements include a consideration of the extent to which the test assesses the skill or knowledge targeted, not other irrelevant characteristics, and the consistency and accuracy of its measurement (i.e. reliability is an essential requirement).

.....

VALUE ADDED

A measure of progress based purely on prior attainment. It takes account of the 'expected' progress of pupils with similar prior attainment and allows comparisons across the key stage for a whole year group (see also contextual value added).

.....

Z-SCORE

Z-scores can highlight how far a pupil's raw score is from the average raw score of test-takers by measuring how many standard deviations below or above the mean raw score their individual score is.



FURTHER RESOURCES

NFER Assessment Hub

The 'go-to' destination for free, digestible guidance on classroom assessment, including:



A series of practical guides

A collection of short-read articles

FAQs answered by assessment experts

Webinars to watch at a time that suits you

Visit www.nfer.ac.uk/assessment-hub to access a host of free support and useful resources to help you build your confidence with classroom assessment.

NFER Tests for years 1-6

Trusted by thousands of schools



Termly assessments to support confident progress monitoring



Trialled with 60,000 pupils to provide robust, comparable data



Supported by a free online tool to generate instant reports



Visit www.nfer.ac.uk/tests to order assessments for your school.

