

Brushing up on assessment

Interpreting information from different assessments

7

One of
ten guides for
school practitioners
and those looking
to brush up on
their assessment
knowledge

Teachers evaluate pupils' attainment based on outcomes from different assessment activities, both formally and informally. Among these activities are tests, questioning, observations and other classroom assessments. In most cases, assessment outcomes complement each other, but occasionally they appear to contradict one another, or to differ from the teacher's expectations based on his or her knowledge of the pupils.

It is advisable to reconcile apparent inconsistencies in order to move learning forward and report accurately to parents and others. This guide offers insight on how to interpret differences in a meaningful way.

Why do assessments give different results?



Since learning is not always a linear process, inconsistencies should be unsurprising. Think of learning, and therefore assessment, as a climbing frame rather than a ladder: pupils may not always continue upwards; they may take a sideways step to reinforce learning, or even take a step down when learning is not fully secure.

Assessment results can differ between subjects when pupils are more secure in their achievement in some subjects than others. Assessments within a subject can also give different results, because pupils can be more secure in some skills or topics than others.

Other reasons for differences can be:

- the mode of the assessment (e.g. pupils may respond differently to practical, written, digital or oral activities)
- the length of the assessment activity
- technical aspects of the assessment (how it is constructed and features such as its reliability)
- the extent of 'scaffolding' or support available
- extraneous factors (such as a pupil's mood, the level of background noise, etc.).

Assessments may give different results for several reasons, and assessment outcomes should be interpreted in context.

Interpretation of assessment results should take factors like these into account. They are the main reason why good assessment is cumulative and needs to include a variety of assessment types over time, as well as outcomes from both **formative** and **summative** assessments.

Drawing conclusions from different assessment outcomes



Remember: all assessment methods have limitations. When inconsistent outcomes arise, use them to prompt questions such as:

- Could an assessment in a different mode have given a different outcome?
- Was there a flaw in the assessment method?
- Did extraneous factors affect the pupil?
- Does the inconsistent result highlight that a pupil has not yet grasped a key concept or skill, despite attaining well in other areas of learning?

Consider all possibilities when faced with discrepancies, and don't attempt to read more into a set of results than is warranted by the information available. Instead use outcomes from different assessments to test your ideas about a pupil's attainment.

Discrepancies can help to build a more complete picture of pupils' attainment, as well as informing ongoing teaching. If assessment outcomes for a pupil are inconsistent, it may simply be that the pupil's knowledge and understanding are well developed in some areas of the curriculum but not yet fully secure in other areas, which continue to need reinforcement. Use such differences to inform your future lesson planning. If inconsistent data suggests poor assessment practice, take steps to improve future assessment activities.

Summarising achievement from different assessments

When summarising achievement, use various assessment outcomes to build a broad picture of attainment. However when compiling data from different assessments, make sure your conclusions are valid by comparing and combining like-for-like.

Be aware that raw scores can only be used to compare a pupil's attainment with others who took the same test, and cannot be compared with a pupil's raw score from different tests, in the same or other subjects, and are therefore not useful for monitoring progress over time. This is because raw scores do not take into account the difficulty of a particular test. For example, NFER's autumn reading tests are based on slightly easier texts than are used later on in the year. If a year 5 child were to complete the autumn, spring and summer year 5 reading tests on the same day at end of the summer term, they would not necessarily get the same raw score across each. Raw scores also don't indicate the total marks attainable in a test. It would therefore be unreasonable to say that a pupil who scored 4, 5, 15, and 32 on four different tests had an average score of 14, if the four tests had different total scores.

Using standardised scores, rather than raw scores, can help you make more accurate judgments of attainment and progress. Standardised scores convert raw scores onto a common scale, enabling test-takers to be compared with a large, nationally representative sample that has taken the test at the same time of year prior to publication. In most cases, the average, nationally standardised score is set at 100, irrespective of the difficulty of the test, so teachers can readily identify whether a test-taker is above or below the national average. This enables comparisons of attainment between different tests, in different subjects and over time.



Using standardised scores, rather than raw scores, can help you make more accurate judgments of attainment and progress.

Managing transition



Inconsistencies can be particularly noticeable at the transition point between key stages. It is common for a pupil's learning to 'dip' over the long summer holiday (the usual time of transition), but even when accounting for this, apparent inconsistencies in achievement may remain.

It is worth remembering that the curriculum content becomes wider with each key stage, and so pupils will need to spend time broadening out that knowledge and consolidating skills in the new context. This does not necessarily mean that the pupil has stopped making progress: it may simply mean that he or she is exploring the 'climbing frame' of learning more fully.

If judgements of a pupil's attainment at the previous key stage do not seem to match those made on transfer to the current key stage, treat the earlier judgements as indicative of the pupil's potential in the wider curriculum of the current key stage.

Found this useful?

This guide is part of NFER Classroom's *'Brushing up on assessment'* series, a collection of free guides providing those looking to build their knowledge with an overview of key elements in assessment.

Sign up to our assessment newsletter (www.nfer.ac.uk/assessment-newsletter) to be the first to receive each guide as soon as it is published, as well as other exclusive assessment-focused content direct to your inbox.

Visit www.nfer.ac.uk/brush-up-assessment to view all the guides in this series.

Looking for assessment materials?

Termly NFER Tests to support attainment and progress monitoring

NFER Tests have been trialled with over 60,000 children to provide standardised scores, enabling you to:

Monitor progress and attainment against national standards.

See how pupils are progressing relative to other pupils their age.

Generate age-related expectation measures in the summer term.

Visit www.nfer.ac.uk/tests to find out more.