Five Myths of Assessment
Removing barriers to focus on assessment principles

Ruth Trundley, Rebecca Cosgrave, Jonathan Gambier, Stefanie Burke and Helen Eversett
Myths of Assessment

Since the removal of levels, schools have been grappling with how to set up effective practices and systems which reflect current expectations. The biggest challenge is to ‘avoid the temptation to recreate levels in a different form’ (John McIntosh, Chair of the Commission on Assessment without Levels: video Sept 2015).

The clear advice from the Commission is to start from principles. It is sometimes hard for teachers to focus on their assessment principles when they are hearing messages which seem to conflict with this advice and instead demand that assessment is shaped in particular ways. In this paper we address five of the most commonly heard ‘myths’ and explain why they should not be diverting schools from their focus on building from principles.

Myth 1: Assessment of every statement within the programmes of study is expected

One of the concerns of the Commission was teacher workload and the final report names unnecessary collecting and recording of data as one of the reasons for an increase in workload:

*The expectation to collect data in efforts to track pupils’ progress towards target levels considerably increased teachers’ workload. The Commission hopes that teachers will now build their confidence in using a range of formative assessment techniques as an integral part of their teaching, without the burden of unnecessary recording and tracking.* (Final Report, Page 14)

Many assessment systems include all of the statements from the programmes of study, leading to an expectation that teachers will assess against every single statement for every child. Recording a judgement against every statement is both time-consuming and unnecessary:

*A school’s assessment system could assess everything students are learning, but then teachers would spend more time assessing than teaching. The important point here is that any assessment system needs to be selective about what gets assessed and what does not, and so the assessment system needs to focus on the ‘big ideas’ in each curriculum area. For example, place value is a central concept in the understanding of our number system. Without a profound understanding of place value, most of mathematics makes little sense. Roman numerals, on the other hand, is not quite so important. As the head teacher or a parent, I would far rather know how a child is doing in terms of their understanding of place value than their knowledge of Roman numerals. You can’t assess everything – be selective.* (Dylan William, Teach Primary)

Being selective, focusing on the ‘big ideas’ is exactly what Tim Oates explained was the thinking behind the removal of levels:

*Assessment should be focused on these key concepts, these key areas of knowledge or skill rather than whether they have achieved a particular level…assessment should therefore focus on has a child understood the key ideas…* (Tim Oates 2015)
The interim standards for teacher assessment at the end of KS1 and KS2 reflect this idea, presenting a selection of ‘key ideas’ for teachers to assess against and this is further supported in the DfE update ‘Five things you need to know about changes to primary assessment’ (February 2016):

*The exemplification should not be seen as a restrictive template on how these judgements should be made and certainly does not require teachers to make checklists of several hundred judgements as has been wrongly claimed. It should go without saying that good teaching is the key to raising standards, not box-ticking.’* (Section 4)

Focussing on the statements within the programmes of study can also mean that important parts of the National Curriculum are ignored and this can distort teaching and learning. For example, in mathematics, the programmes of study do NOT fully reflect the aims of the National Curriculum. These aims are intended to underpin all mathematics:

*The aims of the National Curriculum should be integral to teaching. When developed effectively, they are the key characteristics of good and outstanding practice. Teaching that focuses heavily on covering the listed content, without developing understanding, reasoning and problem solving at the same time is missing the strong drive that the aims represent for improving mathematical education. Such teaching is likely to require improvement.*

(Ofsted 2015)

The aims are expected to underpin all mathematics teaching and shape the mathematical experiences for learners but this will not happen if teachers are focused on assessing the individual statements within the programmes of study. Assessment criteria for maths need to reflect these aims.

There is also no indication of the relative importance of different statements within the National Curriculum, so if each element is assessed they may be assigned the same ‘weight’ even though some have more depth and importance than others. In addition to this, in mathematics, the division of the programmes of study into domains means that key ideas are spread across different domains. Assessing the elements individually will not necessarily indicate that a child has made crucial connections and understood how the maths fits together.

Similarly, within English, connections between different content areas within the curriculum are significant. Whilst it is possible to merely assess individual items, for example the use of fronted adverbials, the way in which a child selects, links and applies each element within a whole piece of writing is fundamentally important.

Assessment should, therefore, focus on ‘big ideas’, the ‘key concepts’ which the children need to understand and schools should be selective about what they record for summative assessments.
Myth 2: Tests provide all the necessary assessment information and need to be used at the end of every teaching sequence or at the end of every half term.

The commitment to assessment which is led by or reliant upon testing rests on the view that it provides an accurate picture of achievement. It separates the learner from any support which would compromise independence or security of learning. The argument goes that if the learner has understood the concept being assessed, then that learning should be able to be reproduced at any time, in any context.

In addition, tests are standardised: a way of measuring pupil achievement against a range of questions or tasks that have been through a process of research to establish their appropriateness to the age of the child and the expectation of their achievement. This standardisation removes the subjectivity of individual teachers’ views on achievement and the setting of questions geared to prove the success of the learning.

Once this view is accepted, then the logical corollary is that a test-focused approach to assessment should not only determine evaluation of achievement at the end of each Key Stage but should also underpin periodic assessment in all other years and the evaluation of pupils’ learning within any year.

Therefore the solution seems simple. In order to evaluate the security of learning, a test approach is more accurate; more broadly representative of the experience of children as they get older and gives confidence that reported results are less subjective and prone to error. Additionally, they are less time consuming and relatively easy to administer providing reports which are easily analysed and shared with a range of audiences.

However, there are a number of issues with relying on tests:

A. ‘Secure’ Learning with Younger Children.

The first issue is the concept of secure learning. Comprehensive anecdotal evidence from teachers suggests that some children struggle to reproduce learning which is secure and widely evidenced in other sources in a test-based situation. Possible causes for this include the emotional pressure of tests on some personalities (see below) or the language used to define a task or question within a test. Some children with well-defined memory and recall skills can also surprise teachers by their achievement in a test, not because the learning is secure or embedded but simply more effectively memorised and recalled. Indeed, the experience of many schools across transition phases seems to indicate that some features of learning demonstrated as ‘secure’ in a test are not committed to longer term memory or fully understood and the children do not retain an ability to apply the learning. Any evaluation of secure learning must encompass an understanding that knowledge and skills have been so learned as to enable reproduction, refinement and application over time—not merely the learning that can be recalled over the short term.
B. The Emotional Impact of Testing Protocols on the Demonstration of Learning:
As noted above, there is growing concern over the emotional impact of continual testing on children’s well-being. The conditions required for tests to be managed to ensure independence run in complete contradiction to learning which is developed and refined on a daily basis. Dialogue and conversation, at the heart of the learning relationship between the children and their teachers, is prohibited. Limiting and sometimes questionable time constraints are imposed. Access to resources and prompts are all suddenly removed. The impact of these approaches can affect children differently but the most vulnerable children emotionally are most likely to be adversely stressed by the experience of testing and therefore less likely to demonstrate their learning capability. Equally, it is self-evident that the impact of these changes will be more apparent where the teaching paradigm shift between exploration and instruction is at its widest, i.e. where the children are of primary age.

C. The Veracity of the Test:
Whilst some tests are standardised, not all undergo rigorous trialling and piloting. Case studies have indicated that results achieved in tests, particularly in years other than 2 and 6, do fluctuate according to the design of the test and to the year group. Historically, the reading assessment for Y5 produced by QCA clearly elevated results, providing Y6 teachers with an inaccurate picture of pupils’ achievement and an incorrect assumption around pupils’ potential outcomes at the end of Y6. Today, the raft of materials available does not always give the same picture of progress when used within schools and across schools. The claim that piloting and trialling tests eradicates these glitches appears to be flawed, raising further questions around the validity of this approach as the sole, or even the most reliable source of evidence for pupil achievement.

D. Breadth and Depth of Learning:
The National Curriculum has as one of its core principles the view that children need to secure ‘mastery’ of learning of concepts appropriate to their age. It has been a concern previously that in the desire to accelerate progress and attainment, teachers have not been able to secure more than a veneer of learning; too thin to facilitate effective recall and application over time. Teachers encouraged by the principle of learning in greater depth have been developing sequences which challenge pupils to explain their learning and to apply their understanding through engaging in problem-solving activity. Already, many exciting learning opportunities are emerging where pupils are choosing the most appropriate resources and equipment to use. They are working collaboratively to challenge each other and to help correct misconceptions and are using the mistakes that they make as powerful drivers for future learning. This activity, though pre-eminent in new approaches to the teaching of mathematics, is also extending to many other areas of subject discipline.

The processes and attitudes required by learners to secure these potential outcomes present a significant challenge to any test developer. In mathematics the test does provide opportunities for pupils to explain and to solve problems but no mathematics ‘test’ can be established which can account for problem-solving activity containing a multiplicity of processes and outcomes. Similarly, mastering concepts also suggests high levels of resilience and perseverance in learning; again difficult to evaluate in a time-constrained assessment.
A further impact of an over-weighted testing approach is the associated narrowing of the curriculum to focus on the content and the procedures of the test. Already, some practice in the development of learning in year 2 in maths is being compromised by teachers denying children the opportunity to use concrete apparatus to support fluency and problem solving in the mistaken notion, that they are preparing children for the ‘experience of the test’. In reading, different elements of the reading curriculum are assessed through teacher assessment, with only some parts of this replicated in tests. Over reliance on reading testing will only deliver a partial view of a child’s reading competence.

E. The Purpose of Assessment:
At the core of the debate around the efficacy of testing lies the central question – what is the purpose of assessment? One purpose is the ability to describe the achievement of children relative to their peers at strategic moments in the school year. Summative assessment of this nature allows measurement of the achievement of pupils and can allow external agencies to evaluate the quality of education being provided by the school.

Additionally, though, both daily and periodic assessment can help teachers to judge how effectively the teaching programme has been in supporting the learning and whether any interventions or changes to the teaching programme may be necessary to tackle underachievement. It is incorrect to assume that testing relates to the former, while teacher assessment facilitates the latter. In reality both features of assessment can be used for either purpose but the judgements made as a result of reflection on either would seem to be more robust and effective when the benefits of both approaches are secured.

F. Periodic Assessment and the Way Forward:
It would seem to be productive, therefore, to maintain a combination of effective daily observation of the work produced by children and the evidence of their learning in class, alongside considered and strategic use of tasks and tests to affirm or challenge the accuracy of this evaluation. Care needs to be taken over the material used for assessment tasks and tests. The problem with these being undertaken at too frequent intervals is that often children are being tested on material they have not actually had the opportunity to learn. Similarly, the danger of making deductions based on learning which is only immediate rather than embedded, explored above, becomes more magnified. Overall, an approach which blends both assessment activities would seem to have the strongest recommendation and legitimacy.
Myth 3: A tracking system underpins good assessment processes

One of the greatest challenges for schools, in moving to a system of assessment without levels, is recognising the difference between tracking and assessment. In the levels-based model, tracking had become, in many cases, the goal of assessment. Teachers completed summative assessments frequently, in order to populate complex and detailed tracking systems that appeared to yield useful evidence of progress and attainment.

Many schools, in moving towards assessment without levels and encouraged by providers of tracking tools, have sought to emulate this model. In many cases tracking systems seek to replicate levels in order to produce data which is easy to track. This is in direct opposition to the statement in the Commission Report on Assessment without Levels (Sept 2015) which states that:

There is a good deal of misunderstanding around the use of the word ‘tracking’ and the Commission has therefore been cautious about using the word in the report. It has become closely associated with measuring progress with levels, in a way that may no longer be appropriate without levels. When evaluating external packages, schools should be aware of this and tread with caution. For example, tracking software, which has been used widely as a tool for measuring progress with levels, cannot, and should not, be adapted to assess understanding of a curriculum that recognises depth and breadth of understanding as of equal value to linear progression... (Final report, page 32)

The Commission Report on Assessment without Levels is clear that schools need to engage in developing models of assessment which:

…directly evaluate(s) pupils’ knowledge and understanding of curriculum requirements (and) helps to create a virtuous circle of teaching and assessment. Teachers assess pupils’ understanding of a topic and identify where there are gaps. This tells the teacher what to focus on in future lessons and prompts the teacher to consider how his or her teaching approach can be adapted to improve pupils’ understanding. (Final Report page16)

Therefore, the primary purpose of formative assessment is to inform future teaching and learning. Summative assessment also contributes to planning for future teaching and learning but additionally provides the school with information they can use to monitor and support pupil progress, attainment and wider outcomes. It is summative assessment which is frequently ‘tracked’. The Commission on Assessment without Levels warns that where tracking systems and the requirements of summative assessment are too great, this can dictate formative assessment in ways which do not support effective teaching and learning. In addition, the report states that:

‘…formative assessments do not have to be measured using the same scale used for summative assessment.’ (Final report, page 15)
This message has been reinforced by a national, statutory assessment model where the outcomes in teacher assessment are not expressed in the same way as outcomes from statutory testing.

The driving force in creating tracking systems for many schools is accountability. Ofsted have been very clear in their guidance to schools about what they would and would not expect to see:

*Inspectors will want to know how schools are assessing whether their pupils are making progress which is appropriate for their age and ability and is sufficiently challenging. Inspectors will gather information from observations in lessons, pupils’ work, discussions with pupils about their understanding and acquisition of knowledge, and the school’s own records. However, Ofsted will not expect any particular data outputs from a school assessment system. (Commission on Assessment without Levels: final report, page 37)*

In this period of change, schools need to be focused on creating assessment policy and practice which:

- Clearly identifies how assessment outcomes will be used, with a view to collecting (and tracking) data only where necessary and only where the data collected is valid and evidence based.
- Is focused on effective formative assessment which is evidence based, valid, useful and draws on a wide variety of assessment strategies.

**Myth 4: Individual children’s books should contain all the evidence of progress and attainment**

The latest Ofsted common inspection framework (August 2015) is putting greater emphasis upon examining the progress of the children currently in the school. Ofsted inspectors will be looking for “progression in knowledge and understanding” in children’s books as part of the assessment information they use to judge a school (DfE 2015 page 21, paragraph 62).

Although a child’s book can be an important source of information about a child’s progress, it is less important for some subjects than for others. Whilst evidence of progress in writing will be mainly in children’s books from across the curriculum, evidence of progress in reading and maths is likely to occur in a variety of places, most powerfully when the children talk about their thinking and understanding. Capturing significant things that children say is therefore an important element of assessment of maths and reading right through the primary age range and should be a major source of evidence.

Evidence in children’s books is particularly powerful if it includes elicitation tasks before sequences of work and application tasks at the end. Many teachers are using elicitation tasks before starting a teaching sequence to ascertain the current understanding of their class. Elicitation tasks can be used to provide evidence of children’s understanding at the start of a sequence, indicating gaps and misconceptions (see references). In maths, children are expected to demonstrate their understanding using different representations and to explain what they don’t understand as well as what they do understand. In writing
children demonstrate what skills they can apply to the given context, showing their understanding of the audience, purpose and form.

At the end of a sequence an application task will provide evidence of new understanding. By comparing the children’s before and after responses to the task it is possible to assess and evidence progress across a teaching sequence. However, it may not always be appropriate for responses from children to be recorded on an individual basis, particularly when considering young children (e.g. in Year 1). Teachers using elicitation tasks to find out what to teach in the next sequence could evidence children’s thinking through observation, such as in role play or small world play, and through the use of probing questions. Examples of the children’s thinking might be found in different places; for example in planning, in a class book with photos and quotes, on the working wall etc.

“There is no intrinsic value in recording formative assessment; what matters is that it is acted on. If it is acted on, there is likely to be other evidence (e.g. in pupil work or lesson plans) to show this.”

(Commission on Assessment without Levels: final report, page 30).

It is useful for work in books to be structured so that changes in a child’s thinking about and understanding of topics being studied are documented and their progress is clear and explicit. If using elicitation/assessment tasks at the beginning and end of sequences, then a child’s book could provide a series of pairs of tasks that could show progress in their thinking and understanding about National Curriculum programmes of study. The work in between these two tasks may show what learning the child has engaged with to make this shift in understanding. If much of this work has been recorded by children on whiteboards, then much of this evidence of thinking and changes in thinking is easily lost. Therefore children’s jotting and thinking needs to be kept; this evidence may or may not be in children’s books, it may be on the working wall, on teachers’ plans, in their assessment notes or in a class book etc.

Evidence of thinking in children’s maths books will not necessarily be neat and well presented as it is not intended for an audience but for the child themselves. In both maths and writing, a book scrutiny that involves the children will reveal more information more quickly than looking at the books alone. For example, the child can be asked ‘Can you show me something in your book that shows you using something you have learned and explain it?’, ‘Can you show me something in your book that your teacher has written that has helped you and explain how it helped you?’ or ‘Can you show me something in your book that you are really proud of and explain why?’.

It is important to evaluate the effect that the marking is having on the child’s learning rather than the quantity of teacher/child’s comments or colour in which it is written. In this way, talking to the child alongside their book provides more evidence of progress than the book can provide on its own and with younger children, who are more likely to have been given oral feedback, the discussion might be more about reflecting on photographs related to their learning.
“Ofsted recognises that marking and feedback to pupils, both written and oral, are important aspects of assessment. However, Ofsted does not expect to see any specific frequency, type or volume of marking and feedback; these are for the school to decide through its assessment policy. Marking and feedback should be consistent with that policy, which may cater for different subjects and different age groups of pupils in different ways, in order to be effective and efficient in promoting learning.”

(Ofsted Handbook, 2015 page 11)

**Myth 5: In maths, children must use formal written methods for calculation**

The National curriculum for mathematics includes the aim that ‘…all pupils become fluent in the fundamentals of mathematics.’ Many people have interpreted this as requiring the fluent use of written methods and written methods only; this is in fact the opposite of the intention of this aim.

Fluency is about more than memorising methods. Russell (2000) writes that fluency includes:

- an understanding of the meaning of the operations and their relationships to each other - for example, the inverse relationship between multiplication and division;
- the knowledge of a large repertoire of number relationships, including the addition and multiplication "facts" as well as other relationships, such as how 4 x 5 is related to 4 x 50;
- a thorough understanding of the base ten number system, how numbers are structured in this system, and how the place value system of numbers behaves in different operations – for example, that 24 + 10 = 34 or 24 x 10 = 240.

All of this is underpinned by the expectation that children will be making decisions based on the numbers and the context. One of the dangers is that an exclusive focus on written methods will in fact limit fluency and make children less flexible in their mathematical thinking.

The Key Stage two tests are often cited as the reason for believing that calculation is all about the written methods, but the tests do not back this up, in fact they do the opposite; success in the tests will require the children to be flexible, use what they know and make decisions about when to calculate mentally and when they need to write something down or use a written method. Full marks are obtained by getting all the answers correct and it is unlikely that a child using a written method for every question will be able to do this in the time available for the tests. For example, in the sample arithmetic paper 6.1 + 0.3 is a simple and quick mental calculation, based on an understanding of the numbers and it would be expected that a child would solve this mentally, using what they know. Recording it as a standard written method will not only take longer, but also will build in the possibility of errors, especially if the children focus on the procedure rather than the numbers involved. The sample arithmetic paper contains four two mark questions and again, the correct answers will give a child all the marks. In theory it is possible to gain one mark for an incorrect answer but the criteria for this make it extremely unlikely, as only one error can be accepted and it cannot be an error related to place value, which is one of the most common errors when using a written calculation for multiplication or division.
One of the key indicators that it is NOT all about written methods is that the interim standard for working at expected at the end of Key Stage two has as its second statement:

*The pupil can calculate mentally, using efficient strategies such as manipulating expressions using commutative and distributive properties to simplify the calculation* (STA, 2015)

A focus on making decisions, understanding additive and multiplicative relationships and ‘using what you know’ will be the best way to support children to become flexible in solving problems and to develop fluency.

**Summary**

As schools continue to develop their own approaches to assessment, the key driver should be the school’s assessment principles. Schools many need to review decisions made to ensure that they fit with these principles; it may mean changing things which do not fit, for example abandoning a tracking system which recreates levels.

Any assessment policy should start from and reflect the school’s principles throughout. Quality teaching will give rise to quality evidence of learning and this should be used to inform future teaching. Using approaches such as elicitation and application tasks will provide vital assessment information to inform planning and also can be used to evidence progress.
References
Babcock LDP maths team have produced two documents ‘Elicitation Tasks in Mathematics’ and ‘Evidencing and keeping track of children’s progress in mathematics’
http://www.babcockeducation.co.uk/ldp/primarymaths - look for the assessment folder
Babcock LDP English team have a similar document ‘Elicitation Tasks in English’
http://www.babcock-education.co.uk/ldp/literacy

Commission on Assessment without levels, (September 2015), Final Report of the Commission on Assessment without Levels

Commission on Assessment without levels, chair John McIntosh (September 2015), video
https://www.youtube.com/watch?v=Wlfv5uy2YUo

DfE (February 2016) Five things you need to know about changes to primary assessment

Oates, Tim (2015) Opening the door to deeper understanding (Cambridge Assessment)


Standards and Testing Agency, DfE (September 2015) Interim teacher assessment frameworks at the end of Key Stage two
