

Report of the NAHT Commission on Assessment

February 2014



Contents

	Page
Foreword by The Lord Sutherland of Houndwood KT	4
Executive summary	5
Summary of recommendations	6
Underpinning principles of assessment	8
Assessment in schools – design checklist	9
Introduction, remit, membership and outline of process	11
The purpose of assessment	12
Current position on levels and testing	13
National curriculum tests and teacher assessment	14
Teacher assessment and moderation	15
Role of accountability and inspection	16
Teacher training	17
The use of information technology (IT)	18
Going forward	18
Annexes	20
Evidence submitted to the Commission	20
Case studies:	
• The Wroxham School, Alison Peacock DBE	20
• Shaftesbury High School, Kerry Sternstein	21
• Cannon Lane Primary School, Reena Keeble DBE	23
History and context of assessment:	
• History of National Curriculum assessment and reporting	24
• History of GCE and GCSEs	29
Assessment practice: an international snapshot	31

NAHT Commission on assessment

Foreword by Lord Stewart Sutherland Chairman's introduction

The decision of the NAHT to set up an independent commission on testing and assessment in schools, which I was asked to chair, is a consequence of the decision of the DfE, following a recommendation by the expert panel for the review of the National Curriculum, to abandon the use of levels and level descriptors in the assessment of school pupils. Two consequences followed: the first was uncertainty amongst many teachers about how they were to carry out the task of assessing pupils' progress across and between school years; the second was a growing realisation of the need, as well as the opportunity, to carry out a thorough review of the role of assessment in schools. I congratulate the NAHT on rising to the latter challenge.

This report is the first stage in this process. In view of the need to offer an approach to tackling the issues which teachers will face in September of this year, this can only be the first stage of the fuller review, which we hope will now engage the profession and relevant government bodies – the DfE, Ofsted and Ofqual. My opening remarks will be focussed upon various headlines, all discussed more fully in the main text.

Those who cannot assess cannot teach.

Assessment is inevitably part of every teaching activity. "How is she getting on?" "Did he understand that?"

Assessment is therefore too important to be the sole preserve of national tests and assessments.

In good education, assessment is of the progress of the whole pupil throughout their educational journey.

Assessment is the means used by good teachers to evaluate that progress and diagnose the needs of the pupil.

True assessment is neither wholly formative, nor wholly summative; it is embedded in the classroom rather than an activity of reflection outside the classroom.

Assessment helps pupils engage more fully in their own development and learning.

A pupil responds better to new challenges if they grasp what is necessary for progress and why.

Assessment helps parents to understand and, as relevant, participate in their children's educational journey.

Quite reasonably, parents want to know how their son or daughter is progressing, and how they can help.

Assessment helps head teachers and governors to plan strategically the use of the resources of the school.

If whole or part classes are not making reasonably expected progress, there could be a variety of causes, and dealing with the uncovered needs may require redeployment of resources in the school.

Assessment of individual pupils and school accountability are interdependent.

One critically important role of assessment is to help appropriate types of the accountability of schools to parents, governors, local authorities and government and tax payers.

Assessment includes externality and objectivity.

This is the main reason for the use of national testing procedures, and also developing the role of in-school, inter-school and external moderation of teacher assessment judgements.

Assessment skills are not sufficiently prioritised in either initial teacher education or continuing professional development.

There is an unjustified assumption at large that assessment is a natural intuitive skill possessed by all.

Assessment will benefit from the fast developing techniques of full pupil profiling which are being enhanced by information technology (IT).

We saw some good examples of schools exploiting this expanding technology for the benefit of all of the above.

Finally, as September approaches,

Don't panic.

There will be a mixed economy in most schools as they see current pupils through the final years of the old system and engage with the new curriculum. Schools are advised to evolve new structures, rather than try to cope with a barren landscape devoid of the old.

Our recommendations will, we hope, help schools in the short term, as well as invite them to engage in the wider debate about the various types and roles of assessment in the medium and longer term.

May I thank all those who submitted written evidence and met with the Commission to discuss

these matters. My fellow commissioners, my colleagues from NAHT and the observers from the DfE, Ofsted and Ofqual, each helped test the evidence which we received and reviewed rigorously, as well as to keep the Chairman on his toes.

Executive summary

The NAHT established its Commission on assessment to focus on finding a way to support schools in determining new assessment arrangements in relation to the curriculum and pupils' learning. The need for this arose following the Secretary of State's decision to remove levels and their associated descriptors from the National Curriculum. In carrying out its task, the Commission was asked to achieve three distinct elements:

- A set of agreed principles for good assessment
- Examples of current best practice in assessment that meet these principles
- Buy-in to the principles by those who hold schools to account.

The Commission, comprising a panel of experienced practitioners, met during the autumn of 2013. The report reflects both the written and oral evidence submitted to the panel, and the subsequent discussions. It contains the Commission's recommendations, a set of principles of good assessment, and a design checklist for a practical assessment framework. It focuses on both the short term implications of the Secretary of State's decision to remove National Curriculum Levels and the medium term consideration of the nature of assessment.

The work of the Commission was, of necessity, urgent because of the timing of the Secretary of State's decision to remove levels from the assessment structure of the National Curriculum and the nature of the changes facing schools with regard to a revised National Curriculum and assessment framework. However, it was felt important at least to begin the work of developing assessment practice for the longer term. The members of the Commission freely acknowledge that there is further work still to do in this area and feel that it would be of benefit for this work to continue.

Summary of recommendations

1. Schools should review their assessment practice against the principles and checklist set out in this report. Staff should be involved in the evaluation of existing practice and the development of a new, rigorous assessment system and procedures to enable the school to promote high quality teaching and learning.
2. All schools should have clear assessment principles and practices to which all staff are committed and which are implemented. These principles should be supported by school governors and accessible to parents, other stakeholders and the wider school community.
3. Assessment should be part of all school development plans and should be reviewed regularly. This review process should involve every school identifying its own learning and development needs for assessment. Schools should allocate specific time and resources for professional development in this area and should monitor how the identified needs are being met.
4. Pupils should be assessed against objective and agreed criteria rather than ranked against each other.
5. Pupil progress and achievement should be communicated in terms of descriptive profiles rather than condensed to numerical summaries (although schools may wish to use numerical data for internal purposes).
6. In respect of the National Curriculum, we believe it is valuable – to aid communication, comparison and benchmarking – for schools to be using consistent criteria for assessment. To this end, we call upon the NAHT to develop and promote a set of model assessment criteria based on the new National Curriculum.
7. Schools should work in collaboration, for example in clusters, to ensure a consistent approach to assessment. Furthermore, excellent practice in assessment should be identified and publicised, with the Department for Education responsible for ensuring that this is undertaken.
8. External moderation is an essential element in producing teacher assessment that is reliable and comparable over time, and all schools should take part in such moderation. Schools should be prepared to submit their assessment to external moderators, who should have the right to provide a written report to the head teacher and governors setting out a judgement on the quality and reliability of assessment in the school, on which the school should act. The Commission is of the view that at least some external moderation should be undertaken by moderators with no vested interest in the outcomes of the school's assessment. This will avoid any conflicts of interest and provide objective scrutiny and broader alignment of standards across schools.
9. Schools should identify a trained assessment lead, who will work with other local leads and nationally accredited assessment experts on moderation activities.
10. Ofsted should articulate clearly how inspectors will take account of assessment practice in making judgements and ensure both guidance and training for inspectors is consistent with this.
11. The Ofsted school inspection framework should explore whether schools have effective assessment systems in place and consider how effectively schools are using pupil assessment information and data to improve learning in the classroom and at key points of transition between key stages and schools.
12. The Department for Education should make a clear and unambiguous statement on the teacher assessment data that schools will be required to report to parents and submit to the Department for Education. Local authorities and other employers should provide similar clarity about requirements in their area of accountability.
13. The education system is entering a period of significant change in curriculum and assessment, where schools will be creating, testing and revising their policies and procedures. The government should make clear how they will take this into consideration when reviewing the way they hold schools accountable as new national assessment arrangements are introduced during 2014/15. Conclusions about trends in performance may not be robust.

- 14.** Further work should be undertaken to improve training for assessment within initial teacher training (ITT), the newly qualified teacher (NQT) induction year and on-going professional development. This will help to build assessment capacity and support a process of continual strengthening of practice within the school system.
- 15.** The Universities' Council for the Education of Teachers (UCET) should build provision in initial teacher training for delivery of the essential assessment knowledge.
- 16.** All those responsible for children's learning should undertake rigorous training in formative, diagnostic and summative assessment, which covers how assessment can be used to support teaching and learning for all pupils, including those with special educational needs. The government should provide support and resources for accredited training for school assessment leads and schools should make assessment training a priority.
- 17.** A number of pilot studies should be undertaken to look at the use of information technology (IT) to support and broaden understanding and application of assessment practice.
- 18.** The use by schools of suitably modified National Curriculum levels as an interim measure in 2014 should be supported by the government. However, schools need to be clear that any use of levels in relation to the new curriculum can only be a temporary arrangement to enable them to develop, implement and embed a robust new framework for assessment. Schools need to be conscious that the new curriculum is not in alignment with the old National Curriculum levels.
- 19.** To assist schools in developing a robust framework and language for assessment, we call upon the NAHT to take the lead in expanding the principles and design checklist contained in this report into a full model assessment policy and procedures, backed by appropriate professional development.
- 20.** Schools should be asked to publish their principles of assessment from September 2014, rather than being required to publish a detailed assessment framework, which instead should be published by 2016. The development of the full framework should be outlined in the school development plan with appropriate milestones that allow the school sufficient time to develop an effective model.
- 21.** A system wide review of assessment should be undertaken. This would help to repair the disjointed nature of assessment through all ages, 2-19.

Underpinning principles for assessment

The principles, in conjunction with the design checklist that follows, will assist schools as they develop their own assessment systems. Schools will be able to review their own processes to ensure that they are underpinned by these principles and, where this is the case, determine whether the assessment system is fit for purpose.

1. Assessment is at the heart of teaching and learning.
 - a. Assessment provides evidence to guide teaching and learning.
 - b. Assessment provides the opportunity for students to demonstrate and review their progress.
2. Assessment is fair.
 - a. Assessment is inclusive of all abilities.
 - b. Assessment is free from bias towards factors that are not relevant to what the assessment intends to address.
3. Assessment is honest.
 - a. Assessment outcomes are used in ways that minimise undesirable effects.
 - b. Assessment outcomes are conveyed in an open, honest and transparent way to assist pupils with their learning.
 - c. Assessment judgements are moderated by experienced professionals to ensure their accuracy.
4. Assessment is ambitious.
 - a. Assessment places achievement in context against nationally standardised criteria and expected standards.
 - b. Assessment embodies, through objective criteria, a pathway of progress and development for every child.
 - c. Assessment objectives set high expectations for learners.
5. Assessment is appropriate.
 - a. The purpose of any assessment process should be clearly stated.
 - b. Conclusions regarding pupil achievement are valid when the assessment method is appropriate (to age, to the task and to the desired feedback information).
5. Assessment should draw on a wide range of evidence to provide a complete picture of student achievement.
- d. Assessment should demand no more procedures or records than are practically required to allow pupils, their parents and teachers to plan future learning.
5. Assessment is consistent.
 - a. Judgements are formed according to common principles.
 - b. The results are readily understandable by third parties.
 - c. A school's results are capable of comparison with other schools, both locally and nationally.
6. Assessment outcomes provide meaningful and understandable information for:
 - a. pupils in developing their learning;
 - b. parents in supporting children with their learning;
 - c. teachers in planning teaching and learning. Assessment must provide information that justifies the time spent;
 - d. school leaders and governors in planning and allocating resources; and
 - e. government and agents of government.
7. Assessment feedback should inspire greater effort and a belief that, through hard work and practice, more can be achieved.

Assessment in schools

Design checklist

These statements provide an evaluation checklist for schools seeking to develop or acquire an assessment system. They could also form the seed of a revised assessment policy; there is certainly value in schools using broadly consistent approaches to assessment.

Our approach to assessment

- Assessment is integral to high quality teaching and learning. It helps us to ensure that our teaching is appropriate and that learners are making expected progress.
- All staff are regularly trained in our approach to assessment.
- We have a senior leader who is responsible for assessment.
- Each pupil is assessed as either 'developing', 'meeting' or 'exceeding' each relevant criterion contained in our expectations for that year. (*note D*)
- Where a pupil is assessed as exceeding the relevant criteria in a subject for that year they will also be assessed against the criteria in that subject for the next year. For those pupils meeting and exceeding the expected standards, we provide more challenging work.
- Assessment judgements are recorded and backed by a body of evidence created using observations, records of work and testing.
- Assessment judgements are moderated by colleagues in school and by colleagues in other schools to make sure our assessments are fair, reliable and valid. (*note E*)

Our method of assessment

- Assessment serves many purposes, but the *main* purpose of assessment in our school is to help teachers, parents and pupils plan their next steps in learning.
- We also use the outcomes of assessment to check and support our teaching standards and help us improve.
- Through working with other schools and using external tests and assessments, we will compare our performance with that of other schools.
- We assess pupils against assessment criteria, which are short, discrete, qualitative and concrete descriptions of what a pupil is expected to know and be able to do.
- Assessment criteria are derived from the school curriculum, which is composed of the National Curriculum and our own local design. (*note A*)
- Assessment criteria for periodic assessment are arranged into a hierarchy, setting out what children are normally expected to have mastered by the end of each year. (*note B*)
- The achievement of each pupil is assessed against all the relevant criteria at appropriate times of the school year. (*note C*)

Our use of assessment

- Teachers use the outcomes of our assessments to summarise and analyse attainment and progress for their pupils and classes.
- Teachers use this data to plan the learning for every pupil to ensure they meet or exceed expectations. Teachers and leaders analyse the data across the school to ensure that pupils identified as vulnerable or at particular risk in this school are making appropriate progress and that all pupils are suitably stretched.
- The information from assessment is communicated to parents and pupils on a termly basis through a structured conversation. Parents and pupils receive rich, qualitative profiles of what has been achieved and indications of what they need to do next.
- We celebrate all achievements across a broad and balanced curriculum, including sport, art and performance, behaviour, and social and emotional development.

Notes and commentary on the design checklist

These notes expand on the statements above with further implications or options.

The types of assessment discussed here are primarily for learning. The information generated is to be used by several different people to plan future approaches to learning. Assessment should not be a bureaucratic exercise for its own sake. The processes should be streamlined to ensure that only those are used that provide information that is useful to teachers, pupils, parents and school leaders.

- A:** There is a task of work to translate the National Curriculum (and any school curricula) into discrete, tangible descriptive statements of attainment – the assessment criteria. As there is little room for meaningful variety, we suggest this job be shared between schools. In fact, NAHT is commissioning a model document.
- B:** The most natural choice of hierarchy for criteria is by school year (certainly the curriculum is usually organised into years and terms for planned delivery). However, children's progress may not fit neatly into school years, so we have chosen the language of a hierarchy of expectations to avoid misunderstandings. Children may be working above or below their school year and we must ensure we value the progress of children with special needs as much as any other group. The use of P scales here is important to ensure appropriate challenge and progression for pupils with SEN.
- C:** We assume that schools will conduct formal assessments more than once a year (and informal assessment will take place continually). A formal assessment at the end of each term, against the year's criteria, is a natural pattern, although some schools will want to do this more frequently. It will take time before schools develop a sense of how many criteria from each year's expectations are normally met in the autumn, spring and summer terms, and this will also vary by subject. Consequently it will also be hard to use this framework by itself for prioritising intervention in the first few years of use.

For some years to come, it will be hard to make predictions from outcomes of these assessments

to the results in KS2 tests. Such data may emerge over time, although there are question marks over how reliable predictions may be if schools are using incompatible approaches and applying differing standards of performance and therefore cannot pool data to form large samples.

- D:** There is a need to record a pupil's attainment against each applicable assessment criterion. The criteria themselves can be combined to provide the qualitative statement of a pupil's achievements, although teachers and schools may need a quantitative summary. Few schools appear to favour a pure 'binary' approach of yes/no. The most popular choice seems to be a three phase judgement of working towards (or emerging, developing), meeting (or mastered, confident, secure, expected) and exceeded. Where a student has exceeded a criterion, it may make sense to assess them also against the criteria for the next year.

These recorded judgements can be translated into numbers, which can then be analysed and used for prioritising. Traffic lighting is a popular method for monitoring. The most obvious method to generate a 'colour' or status is to count the proportion of the relevant year's criteria that have been met at that point in time. At this stage, it is not possible to say what proportions would be cause for concern or celebration at a particular time of the year – although presumably you would expect to have mastered all applicable criteria to be green at the end of the year.

The method of 'fitting' a student to a criterion must be consistent to draw comparisons between groups. If the criteria are discrete, concrete and precise, this will remove some ambiguity. If a school is using a three phase judgment, one would expect the middle 'meeting' to be based on mastery.

- E:** The exact form of moderation will vary from school to school and from subject to subject. The majority of moderation (in schools large enough to support it) will be internal but all schools should undertake a proportion of external moderation each year, working with partner schools and local agencies. It is also good practice to invite external agencies with no connection to the local group of schools to verify practice from time to time.

Introduction

In the summer of 2013 the government announced the end of the official use of National Curriculum levels for assessment, following a recommendation from the expert group on National Curriculum Review. This caused concern across the profession and gave rise to such questions as how inspectors would react to multiple different assessment systems in place in schools, how progress would be demonstrated and judged, and how attainment would be measured. 'Levels' had become the accepted language both of pupil attainment and progress and the prospect of the removal of this language caused widespread consternation.

The government has a stated policy of freedom and autonomy for school leaders. NAHT believes strongly that freedom need not mean fragmentation and, if the government wants to transfer ownership of assessment to the profession, then the profession should take that ownership and design a proper replacement. The removal of levels provides an opportunity for the government, its agencies and, most importantly, the profession itself to enhance the professionalism of teachers in the development and use of assessment. In furtherance of this aim, NAHT decided, therefore, to establish an independent commission on 'assessment without levels' to consider what lay behind good assessment and to look for examples of good practice already in place or developing in schools.

Remit of the Commission

The Commission was asked to achieve three distinct elements, with their associated outcomes:

A set of agreed principles for good assessment;

Examples of current best practice in assessment that meet these principles; and,

Buy-in to the principles by those who hold schools to account.

The remit did not extend to KS2 tests, floor standards and other related issues of formal accountability. Rather, the focus was on assessment for learning within school. During its considerations, the Commission decided it would be helpful to outline not only high level principles of assessment

but also a more practically-based design checklist to assist schools in the short term as they review their assessment framework.

The Commission met during the autumn of 2013, with a panel of experienced practitioners, including official observers from Ofsted, Ofqual and the Department for Education (DfE). A formal invitation was issued to stakeholders to submit evidence for consideration and the Commission received both written and oral evidence. The report of the Commission contains a series of recommendations covering the elements listed above.

We hope that these recommendations and principles will support consistency in assessment without constraining freedom. Above all, they should give schools confidence that, if they invest in developing approaches to assessment that accord with these recommendations, principles and checklists, inspectors and officials will give due credence to those systems.

Membership

Membership of the Commission ranged across the education profession. The panel was chaired by former HMCI of schools **Lord Stewart Sutherland** and its members were:

Leora Cruddas

Director of Policy, ASCL

Tony Draper

Head Teacher, Water Hall School

John Dunford

John Dunford Consulting

Hilary Emery

CEO, National Children's Bureau

Sam Freedman

Director of Research, TeachFirst

Russell Hobby

General Secretary, NAHT

Bernadette Hunter

NAHT National President 2013-14

Kathryn James

Director of Education, NAHT

Steve Kirkpatrick

Deputy Head, Willow Tree Primary

Dame Alison Peacock

Head Teacher, Wroxham Primary

Tim Sherriff

Head Teacher, Westfield School

Kerry Sternstein

Deputy Head, Shaftesbury School

Prof Gordon Stobart

Institute of Education

Observers from Ofqual, Ofsted and DfE

Supported by **Mick Walker** and **Alex Rowley**

Outline of the process

In carrying out its work, the Commission sought evidence for consideration from as wide a range of stakeholders as possible. With this in mind, a public call for evidence was extended and individual invitations were addressed to specific stakeholders. Those willing to present evidence (both oral and written) were asked to address the following questions:

- What are the purposes of assessment?
- Who benefits from assessment?
- What are the elements of good assessment practice?

- Is a universal system of assessment necessary to measure pupil progress and attainment?
- What aspects of learning should be assessed and how?
- What forms of assessment are appropriate for use at the following ages?
 - 0-4yrs (early years)
 - 5-7yrs (KS1)
 - 7-11yrs (KS2)
 - 11-14yrs (KS3)
 - 14-16yrs (KS4)
- What should be the outcomes of an effective assessment system?
- What quality assurance mechanisms are needed to ensure robustness and reliability of assessment?
- What role should assessment play in formal inspection?
- What other areas of assessment should be considered by the Commission?

The Commission met to consider the written evidence received and also took oral evidence from a wide range of sources. We would like to extend our gratitude to the individuals and organisations who so freely gave their time and expertise in submitting their views. Details of the evidence received are given in the annex.

The Commission's view on the purpose of assessment

Before looking in detail at the evidence submitted, the Commission spent time reflecting on the nature of assessment and its purposes. In an educational context, the term 'assessment' is used to denote a range of measurement functions for formative, diagnostic and summative uses. The data derived from such assessments are used to:

- inform pupils, parents and others about the performance of individual pupils and to inform teaching and learning;
- hold schools accountable for the attainment and progress of their pupils;
- enable benchmarking between schools as well as monitoring performance both locally and nationally; and
- award qualifications such as GCSE and GCE A level.

Assessment has many forms, including different types of testing and individual teacher assessment through observations of pupils in class. It is easy to conflate assessment with testing, but testing is merely one method of assessment. The Commission's task was to look into a system that, from September 2014, will be based on a new National Curriculum for all maintained schools that would no longer be underpinned by nationally determined performance levels. Therefore, the Commission's focus was on finding a way to support schools in determining new assessment arrangements in relation to the curriculum and pupils' learning.

Teachers assess pupils' progress on an ongoing basis in the classroom, determining what is being learned, what pupils know, understand and can

do and what they need to do next to progress. Arguably, this formative and/or diagnostic assessment is the basis of nearly all teaching and, without it, what happens in the classroom runs the risk of being directionless and ineffective.

The use of level descriptions came into being in the early 1990s, shortly after the introduction of the National Curriculum, and was designed to enable consistency and uniformity in describing what an individual pupil had attained and thereby the progress being made. This has developed into a common language, very often shortened to a numerical value, used by schools and others as a shorthand to summarise progress and attainment. A numerical summary can be useful to schools for the purpose of analysis and tracking over large numbers of pupils but it misses the richness of a more rounded description of achievement. (The use of P

scales for pupils with special educational needs has been developed and continues to be important to ensure appropriate challenge for children with SEN.)

The original level descriptors have evolved in line with various revisions of the National Curriculum in terms of their definition and the uses to which they have been put. As time has progressed, the descriptors have been assimilated into the common language. The important narrative behind the headline number was often lost through over-simplification. Thus, the assumption was made that we mean exactly the same thing when we refer to a pupil achieving a certain level: level 4, for example. Of course, this is not necessarily the case. Using level descriptor shorthand also fails to address whether the pupil is consistently achieving that level, achieving it most of the time or on occasion or, indeed, only in some aspects. It also fails to show a pupil's particular strengths and weaknesses.

Current position on levels and testing

The decision of the Secretary of State to remove the official use of levels and level descriptors, although arising from the report of the National Curriculum review expert panel, was unexpected and caused concern within the profession. Although the government stated that levels were not clearly understood by parents and other stakeholders, this was disputed by teachers and other education professionals. Many said that this current generation of parents had grown up with such a system of communicating attainment and progress and, as such, had an adequate grasp of what was meant. It became clear to the Commission that there needed to be two points of focus for its work: the **short term** implications of the Secretary of State's decision – what schools should do in September 2014 – and the **medium term** consideration of the nature of assessment more generally.

As well as measuring pupil progress and defining attainment in National Curriculum terms, levels had also been developed further by the profession to address the issue of progress within schools as an element of the accountability system. The original National Curriculum levels 1, 2, 3, etc. were further refined to 2c, 2b, 2a and so on, with the continued

expectation that pupils would make two levels of progress over a key stage. Thus the expected two levels of progress at certain key stages were looked at in terms of six sub-levels. This development of sub-levels was, in the main, profession-led and initially not widely used by official sources. However, more recently, when defining the term, 'secondary ready' in the DfE primary accountability consultation, the government alluded to this being equivalent to level 4b. Given that the expected level of attainment for the majority of pupils at the end of key stage 2 had been level 4, and this notion of 4b was raised after the announcement that levels were to be removed, this only added to the general confusion and concern.

The majority of those giving evidence to the Commission highlighted that level descriptors and National Curriculum levels, whatever their other faults, had given the profession a common tool to communicate with each other and with stakeholders. Others commented that whatever replaces levels needs also to retain a common language or, if a number of different assessment systems emerge, there would still need to be an element of compatibility or common understanding.

National curriculum tests and teacher assessment

Testing is often seen as reliable, definitive and objective. However, the Commission found this far too simplistic a view. There is clearly a value in using tests, both internal and external, and most schools use them internally on a regular basis. However, assessment is not an exact science. Tests are in effect a snapshot of what a pupil can do on that day at that particular time in a specific sample of the curriculum and may or may not be an accurate measure of a pupil's attainment over a wider period; tests are effective for assessments of certain types of knowledge and less effective for others. In a more reliable system, tests should be used to inform and be part of the teacher's wider assessment of pupils' progress and attainment.

The Commission's view that too great a reliance is being put by government on external tests, particularly for school accountability purposes, was widely supported by the evidence submitted to the review. This over-reliance has led to distortion in curriculum emphasis and accusations of 'teaching to the test'. If too much weight is attached to *any* form of assessment it is likely to lead to such perverse incentives – not only teaching to the test, for example, but also inflated teacher assessment because of performance management issues, or deflated assessment to enhance the measurement of later progress. Although assessment is frequently used as the basis of accountability, the very nature of accountability influences the results of assessment, which in turn constrains what forms of assessment may be used. The Commission heard how too restrictive an assessment system can lead to the narrowing of the curriculum. Teachers need the opportunity to develop trust and confidence in their own and colleagues' ability to assess pupils accurately. We need a more coherent, rounded approach to assessment overall and an accountability system that does not raise the stakes and distort outcomes unnecessarily in any one area. This highlights the challenges of using assessment for multiple purposes, because what may suit accountability may damage assessment for learning and the accuracy of judgements.

The Commission's view is clear: the link between curriculum, assessment and pedagogy must not be

ignored. An effective assessment system built around the curriculum will inform classroom practice and help improve teaching and learning. Whilst we note that the revised curriculum has not been dictated to by an assessment framework, curriculum and assessment should be developed in tandem. The Commission heard compelling arguments about the adverse effects of levels caused by the labelling of pupils and the oversimplification of numerical measures. This underpinned the government's decision to remove levels. Ironically the revised curriculum is presented in a model of year-by-year progress with planned descriptors of performance at the end of each key stage. Taken alongside the lack of timely guidance and exemplification, this only exacerbates the demands on individual schools for September 2014. The Commission therefore believes that, so far as schools are following the National Curriculum, it is valuable – to aid communication and comparison – for schools to be using consistent criteria for assessment. The Commission also recommends that the profession acts to create a set of model assessment criteria, based on the new National Curriculum, that demonstrate pupil progress.

- 1. Schools should review their assessment practice against the principles and checklist set out in this report. All staff should be involved in the evaluation of existing practice and the development of a new, rigorous assessment system and procedures to enable the school to promote high quality teaching and learning.**
- 2. All schools should have clear assessment principles and practices to which all staff are committed and which are implemented. These principles should be supported by school governors and accessible to parents, other stakeholders and the wider school community.**
- 3. Assessment should be part of all school development plans and should be reviewed regularly. This review process should involve every school identifying its own learning and development needs for assessment. Schools should allocate specific time and resources for professional development in this area and should monitor how the identified needs are being met.**

4. **Pupils should be assessed against objective criteria rather than ranked against each other.**
5. **Pupil progress and achievements should be communicated in terms of descriptive profiles rather than condensed to numerical summaries (although schools may wish to use numerical data for internal purposes).**
6. **In respect of the National Curriculum, we believe it is valuable – to aid communication and comparison – for schools to be using consistent criteria for assessment. To this end, we call upon the NAHT to develop and promote a set of model assessment criteria based on the new National Curriculum.**

Teacher assessment and moderation

The Commission heard from the majority of those submitting evidence that there was a lack of trust in teacher assessment at the present time. All saw the need to address this urgently; teacher assessment was seen as a vital component of a balanced and reliable assessment system. The unanimous view was that the necessary improvement in teacher assessment would not be achieved by increasing external testing; that, in fact, this risks further deskilling of teachers in broader assessment methods. Rather, much of the evidence was aligned in the view that a general move to external testing over the years had eroded both public and professional confidence in the outcomes of internal assessment.

Concerns were raised about initial teacher training and how it dealt with assessment, about the lack of continuing professional development for teachers in the area of assessment, about developing high quality professional dialogue and, significantly, about the need for effective moderation.

To enable development of teacher expertise, the Commission is of the view that there needs to be a more widespread climate of trust. Professional development of teacher skills will need time. Although there is a general belief that classroom assessment practice has improved in assessment for learning, for example, this needs to be balanced with the lack of trust exhibited by the profession itself – junior schools often report that infant schools' assessments of their pupils are over-inflated, secondary schools argue that they need to test pupils on arrival because primary assessments, including national tests, cannot be relied upon. In part, this lack of trust is due to a lack of consistency and in part to the perverse incentives resulting from a high stakes accountability model.

Trust, of course, must be earned as well as given. Schools must take professional accountability and pride in the integrity and accuracy of their assessments. The level of trust can be extended further by a system in which moderators have the ability and professional standing to contest assessment judgements made at school level where necessary.

A recurring theme in the evidence presented to the Commission was the need for 'externality' in terms of assessment. It was clear that, for any assessment system to be trusted, there had to be a reliable form of external checking. For teacher assessment, this would take the form of external moderation. Examples were given to the Commission of in-school and cross-school moderation which went a long way to delivering on this aim. However, if the moderation process is to be robust, it is important to have a degree of judgement and accountability which was not always present in the examples given.

It is important to distinguish between professional dialogue, standardisation and moderation. Each are relevant and important elements of a robust system, with professional dialogue playing a huge part in teacher development. Moderation should bring with it some 'teeth', some judgement and a requirement to take note. The moderator must have the right, indeed, the responsibility to say that the assessment level is too low, too high or correct, with an expectation that this judgement will be acted upon.

There is a worrying lack of trust in individual teacher-based assessment, which emanates from within the profession itself. For example, the

Commission heard how secondary schools were likely to test pupils as they came into year 7 rather than trust the KS2 assessments. This was generally seen as a problem caused by the nature of the accountability system rather than any underlying lack of ability within the profession. However, there was a clear acknowledgement of the need to develop further teachers' skills in this area.

It was generally appreciated in the evidence given to the Commission that nationally standardised tests have an important part to play in supporting and validating teacher judgements but public trust can be increased by developing the use of effective moderation practices. This can be increased further by improving assessment training in initial teacher training and continuing to develop CPD, including encouraging professional dialogue within and across schools. As previously stated, further national tests were seen generally as unnecessary. Indeed, it was argued by a significant number of those submitting evidence that the current test regime should be evaluated against its intended purposes. Some members of the Commission were of the view that it would be more appropriate to use national sampling to review the performance of all schools at a national level, with improved, moderated teacher assessment being used to hold schools accountable on an individual basis. Other members of the Commission foresaw a continued role for universal testing. Any such system must provide a national picture of performance standards, provide local accountability

and improve the quality of assessment which, most importantly, would improve teaching and learning.

- 7. Schools should work in collaboration, for example in clusters, to ensure a consistent approach to assessment. Furthermore, excellent practice in assessment should be identified and publicised, with the Department for Education responsible for ensuring that this is undertaken.**
- 8. External moderation is an essential element in producing teacher assessment that is reliable and comparable over time, and all schools should take part in such moderation. Schools should be prepared to submit their assessment to external moderators, who should have the right to provide a written report to the head and governors setting out a judgement on the quality and reliability of assessment in the school, on which the school should act. The Commission is of the view that at least some external moderation should be undertaken by moderators with no vested interest in the outcomes of the school's assessment. This will avoid conflicts of interest and provide objective scrutiny and broader alignment of standards across schools.**
- 9. Schools should identify a trained assessment lead, who will work with other local leads and nationally accredited assessment experts on moderation activities.**

Role of accountability and inspection

The Commission considered evidence that commented on the potential for tests and the current accountability system to distort and corrupt the curriculum and the diagnosis of pupil ability and progress. By their very nature, tests sample a relatively small proportion of the curriculum. Indeed, the tests for National Curriculum English, for example, completely omit such essential elements of the subject as speaking and listening. Therefore, when tests play such a central role in the accountability system, it is hardly surprising that this has, in some ways, narrowed

and limited the curriculum unduly. Because the national measurement of progress and attainment is tied closely to accountability, this raises the stakes in terms of the elements which will be tested. This in turn can result in other elements of ability being seen as less important, leading to an overall misdiagnosis of the pupil's actual ability. This can mislead pupils and their parents. There is a concern that only that which can be tested is valued and this is having a negative and restrictive impact on what is taught.

Whilst there was undoubted concern about the current accountability system and its effect on assessment, the Commission was clear that there was a willingness from the profession to be held accountable. In fact, there was considerable support for a more effective form of accountability, one that went beyond the simplistic and potentially damaging snapshot and, instead, looked to get a more rounded picture and, in so doing, recognised teachers as valued professionals. The Commission believes that it is not unreasonable to hold schools accountable for the quality of their assessment practice as well as their assessment results, with head teachers also appraising teachers on their use of assessment.

There was strong concern expressed about the requirement for schools to publish their detailed curriculum and assessment framework in September 2014. A significant number of those submitting evidence stressed the importance of Ofsted making absolutely clear to schools what they expect of schools' assessment processes. There also needs to be absolute clarity from the DfE about expectations of teacher assessment and moderation and what data schools will be required to submit. When these elements are in place, schools would then need time to develop their own assessment and reporting systems. It would therefore be helpful if, rather than asking schools to publish a detailed assessment framework for September 2014, they were asked to publish their *principles* of assessment, with further

detail being published when there had been time to develop and embed robust assessment and reporting systems.

10. Ofsted should articulate clearly how inspectors will take account of assessment practice in making judgements and ensure both guidance and training for inspectors is consistent with this.

11. The Ofsted school inspection framework should explore whether schools have effective assessment systems in place and consider how effectively schools are using pupil assessment information and data to improve learning in the classroom and at key points of transition between key stages and schools.

12. The Department for Education should make a clear and unambiguous statement on the teacher assessment data that schools will be required to report to parents and submit to the Department for Education. Local authorities and other employers should provide similar clarity about requirements in their area of accountability.

13. The education system is entering a period of significant change in curriculum and assessment, where schools will be creating, testing and revising their policies and procedures. The government should make clear how they will take this into consideration when reviewing the way they hold schools accountable during 2014/15. Conclusions about trends in performance may not be robust.

Teacher training

Evidence heard by the Commission was consistent in the view that, in terms of assessment, teacher training was not of a sufficiently high or rigorous standard. This applied across the board, from initial teacher training through to on-going professional development. With any change to the system, there needs to be sufficient support and development for teachers to allow them to adjust and adapt to the change. All teachers are not automatically equipped to assess, even though there is an apparent assumption that this is the case. They need practical training in assessment methodology

and practice and an ongoing programme of CPD. Some also commented that it could be useful for the teaching standards to reflect further assessment knowledge, skills and understanding.

Although the awareness of newly qualified teachers in relation to accountability measures appears to have increased, the same cannot be said for their awareness of assessment practice. In part, this is due to some deficiencies in initial teacher training, as well as the inability or unwillingness of schools to be flexible in their approaches when working with trainee teachers. The latter was highlighted as one way in

which early development of experience and expertise in newly qualified teachers was being stifled.

For more experienced teachers, a comprehensive programme of CPD is required to improve the overall quality, thereby providing greater validity and increased public confidence. A national standard in assessment practice for teachers would be a useful addition. The Commission also favoured the approach of having a lead assessor to work with each school or possibly a group of schools, helping to embed good practice across the profession.

14. Further work should be undertaken to improve training for assessment within ITT, the NQT induction year and on-going professional development. This will help to build assessment capacity and support a process of

continual strengthening of practice within the school system.

15. The Universities Council for the Education of Teachers (UCET) should build provision in initial teacher training for delivery of the essential assessment knowledge.

16. All those responsible for children's learning should undertake rigorous training in formative, diagnostic and summative assessment, which covers how assessment can be used to support teaching and learning for all pupils, including those with special educational needs. The government should provide support and resources for accredited training for school assessment leads and schools should make assessment training a priority.

The use of information technology (IT)

The Commission heard evidence about the desirability of improving the use of technology in assessment and the communication of assessment outcomes. For example, GL Assessment and IAPS both highlighted the benefits of adaptive, online testing in producing more granulated outcomes than traditional testing. The development of e-profiles brings with it the potential for parental access at any time and also ease of parental input and support. A particular strength of IT is that it can be used to track progress as well as collect and report data and other assessment information.

Additional uses of technology include using the internet to facilitate professional dialogue and extending professional communities beyond the traditional and sometimes limited local networks. The panel heard from Streetly School how it is developing links with its feeder schools through anonymous sharing of pupils' work as part of a moderation process.

17. A number of pilot studies should be undertaken to look at the use of IT to support and broaden understanding and application of assessment practice.

Going forward

The Commission set itself two broad aims: the first was to consider how schools could be helped in the short term and to make such recommendations as would be possible for schools to implement within the timescale of September 2014. The second aim was to consider in more detail what would be necessary in assessment terms going forward.

Short term

Given the immediacy of the issues, with schools facing the removal of levels and level descriptors from September 2014, the Commission's recommendations are of necessity interim in character. However, it was recognised that in no way should these disagree with the principles outlined in the report, as these should form the

basis of future developments. It is also important that the short term recommendations are practicable and accessible to the profession.

It is likely that, in September, schools will need to base their practice on the skills and methods currently in place to satisfy the need for good diagnosis of learning needs and appropriate teaching response. However, it is important to recognise that, even if schools decide to continue with some form of levels, the new National Curriculum does not align to the existing levels and level descriptors and this alignment is a piece of work that needs to be undertaken now.

The Commission is concerned that, currently, there is a requirement for schools to publish their detailed curriculum and assessment frameworks by September 2014. This seems unduly hasty and does not give schools sufficient time to develop fully their assessment systems. Rather than expect a fully fledged, deeply embedded assessment system in all schools, it would be more sensible to ask schools to publish their assessment principles in September 2014. It would then be possible to build in a plan to develop a detailed assessment framework, with the full system being in place and published by no later than September 2016.

18. The use of suitably modified National Curriculum levels as an interim measure in 2014 should be supported by the government. However, schools should be clear that any use of National Curriculum levels in relation to the new curriculum can only be considered a temporary arrangement to enable schools to develop, implement and embed a robust framework for assessment. In doing so, schools need to be conscious that the new curriculum is not in alignment with the old National Curriculum levels.

19. To assist schools in developing a robust framework for assessment, we call upon the NAHT to take the lead in expanding the principles and design checklist contained in this report into a full model assessment policy and procedures, backed by appropriate professional development.

20. Schools should be asked to publish their principles of assessment from September 2014, rather than being required to publish a detailed assessment framework, which instead should be published by 2016. The development of the full framework should be outlined in the school development plan with appropriate milestones that allow the school sufficient time to develop an effective model.

Medium term

Much of the evidence given to the Commission highlighted the interest and concern that exists amongst the education profession in improving the assessment process as a whole. There is clearly a need and, importantly, a desire to engage in a much broader examination of the nature of assessment and the varying forms and outcomes which it can have.

To build on this willingness to engage, the Commission believes strongly that further study should be made of the whole system of assessment. Rather than the disjointed picture currently in place, there is a need for a logical, mutually, compatible assessment process that covers the whole of a child's education through whatever educational setting he/she attends.

21. A system wide review of assessment should be undertaken. This would help to repair the disjointed nature of assessment through all ages, 2-19.

Annexes

Evidence submitted to the Commission

Achievement for All	Independent Association of Preparatory Schools	Shaftesbury High School
Assessment Reform Group and British Educational Research Association	Professor John White, Institute of Education	Southampton Local Authority and Teaching Alliance
Association of School and College Leaders	Liz Twist, National Foundation for Educational Research	The Streetly School
Association of Teachers and Lecturers	Mayflower Primary School	Tim Oates, Cambridge Assessment
Corporation Road Primary School	National Children's Bureau and Council for Disabled Children	Universities' Council for the Education of Teachers
Dr Des Hewitt, University of Derby	Oxford University Press	Voice
Geographical Association	Pearson	The Wroxham School
GL Assessment	Professor Peter Tymms, Durham University	
Goddard Park Community Primary School	Ridgewell Primary School	

Case studies

Alison Peacock DBE, Headteacher, The Wroxham School

At The Wroxham School, a primary school with nursery in Hertfordshire, we have not talked to children and parents about 'levels' for the last ten years. Within that time the school has achieved three consecutive Ofsted 'outstanding' judgements and attainment has remained high. The secret of our success has been to create a 'listening' culture where children and adults know that they are trusted and valued. Dialogue about how to support each child and their family has remained the central discourse. Within an environment where everyone believes in the importance of learning, tests and grades take care of themselves.

Life without levels leaves space for the highest quality curriculum, opportunities to tailor learning to the needs of the individual and an environment of excitement and ambition. Children at Wroxham do not rank their performance against their peers but work happily with a wide range of learning partners in each class. They are encouraged to challenge themselves and to make good choices about the complexity and difficulty of a range of tasks. They are not given targets or grades but there is a strong focus on formative feedback. There are no ability based groups and children are offered choice about joining additional sessions if they feel they could benefit. For example, challenging maths lessons are offered by a member of staff several times a week for the older children. If they feel that they can thrive in a fast-paced lesson children have

the option of joining the group. The emphasis is on self knowledge and self regulation that offers constant opportunity for improvement.

Families are very well informed about their children's achievements and progress as this is evidenced in the manner in which each child can talk about his learning and next steps for improvement. The quality of work produced provides excellent opportunities for children and families to witness evidence of progression and increasing skill. This learning is not confined to written work and calculation, but is also evidenced through areas such as sports achievement, musicality and scientific problem solving. End-of-year reports are written by the children from year one to year six and form an electronic dialogue with the teaching team. Photographs illustrating the breadth of the curriculum are uploaded to these documents by the children. The youngest children are supported in the composition of their report comments by older children working with them in pairs.

Learning Review Days and Family Consultations provide families with an opportunity twice a year to review progress in a detailed manner. In years 5 and 6 these meetings are held in my office. As headteacher, I attend every meeting for every year 5 and year 6 child. The meetings are led by the child who prepares a powerpoint presentation for their family and teachers providing feedback about their current challenges and successes. Conversation then ensues with the child at the centre, upon how the adults can support next steps in learning. This is a highly rigorous process, as it means that all participants are committed to supporting the highest possible ambition for that child's success.

In the summer term we organise a whole school open learning event where families arrive to take part in a wide range of activities in classrooms throughout the school and in the grounds. We provide a simple A5 passport for children so that they can collect stamps in each room where they take part in learning with their family during the evening. Feedback from these sessions is very positive.

Families are invited into school on a regular basis. This enables shared understanding of the ways in which we teach and provides opportunities

for children to impress their families with their knowledge and understanding. As a teaching school, Wroxham welcomes many visitors. Our children are very proud, informed guides and are keen to explain about our system of choices in lessons. If you are interested in hearing the children talk about this way of working without levels please visit www.wroxham.net to see films and blogs about day to day learning.

As a head teacher, it is my belief that our experience of rejecting numbers as a proxy for ability has enabled us to focus on the really important work of finding a way through for every single child to learn. We have kept data tracking sheets as a management resource purely as a means of ensuring that no child slips through the net (and to provide a record for Ofsted when they inspect). The crucial difference at Wroxham is that the learning comes first and assessment is used to inform the process but not to provide a label or measure. The outcomes at our school speak for themselves and we are optimistic that others may be interested to join us on this journey now that within England we have a unique opportunity to assess beyond levels. Let's take the opportunity we have been given to change the emphasis in schools to one of 'I can't do it ...yet' rather than 'I can't do it because I'm only level three'.

**Kerry Sternstein, Deputy Headteacher,
The Shaftesbury School**

Shaftesbury High School is a special needs high school in Harrow catering for young people ages 11-18 with complex learning, social, emotional and behavioural needs. Young people whose abilities range from SLD to mainstream (GCSE). We are a specialist school for pastoral care, a national support school and have been consistently outstanding in the last two Ofsted inspections.

We believe that assessment and learning are part of the same continuous process, underpinned in an atmosphere of mentoring and support. The process is designed so assessment and learning inform each other throughout the year on a daily, weekly and termly basis.

Pupils' achievement and progress is monitored by continuous setting and reviewing of targets -

targets not determined to level or judge, but to guide and reinforce. Pupils set their own targets for learning and development with their form tutors at the start of every term. These are reviewed formally in the same way towards the end of the term and informally at the end of each week with the teaching assistant. The meetings with the form tutor are held in private and 1:1, something pupils and staff alike find a very positive experience. Pupils and staff together decide whether they have met the target (green), are still working on the targets (amber) or have not met it (red), perhaps suggesting changes are required. If a target has been met or is no longer appropriate, it can be changed at any time.

All targets at Shaftesbury can be modified, extended or revisited. The targets are SMARTER, positive and meaningful. They may be academic or behavioural or a mix, either way they can be applied across the curriculum. Academic targets are based on descriptors and 'I can' statements, not on NC levels, although they are relevant to them. Staff also set subject targets (ILOs) with and for pupils. These are continuously reviewed throughout the year and used in every subject area. Some are taken from attainment level statements, some may refer to a specific skill needed for accreditation. All are designed to recognise areas of progress, address areas of concern and offer the next steps on each student's learning path. These are also discussed during the three parent/teacher consultation evenings held across the year and at annual reviews. This process is rigorous and consistent to ensure all stakeholders share in decisions about pupils next steps.

Annual review meetings are a compulsory requirement for all statemented pupils. This is why pupils are at the centre of our annual reviews. They bring examples of work from their record of achievement, on paper and in a digital format. They also choose their favourite photographs and talk about the occasions when these were taken. They complete self assessments, choosing from a range of options, smiley faces, written text, multiple choice questions and verbal feedback. These feed into an annual review front cover summary sheet, designed to be accessible for the student and their

parent-carers. Many of our parent-carers have significant needs of their own, and we want all the information that we share to be part of a proactive, mutually nurturing dialogue. Staff have worked long and hard to defuse the anxieties around data and assessment, to ensure they are part of a truly inclusive package.

In the upper school and the 6th form, the pupils have visual annual reviews. The pupil is at the centre of this, directing what is being recorded, under a number of headings such as: home, school, college, my friends, important people, next steps. Drawings or symbols are used to illustrate the students' comments and they sign it to assure ownership. These annual reviews have been greeted with great positivity both from students and parent/carers who feel they can understand what is being said and what is being discussed and that the young person is really part of it.

The 6th form students also collect evidence of work and experiences throughout their two year course. This is often shared during annual reviews but is also stored on pen-drives and presented when they leave. This provides an excellent, detailed and accurate record of attainment in a visual and creative format for them to take to college or the next stages of their education and development. There are also plans to produce hardback format versions for each student, and a 'greatest hits' legacy version which all can contribute to with the aim of inspiring their successors.

All reports are written in advance of annual reviews. The reports aim to acknowledge pupils' strengths, development and progress, discuss their ILOs and comment on attainment. These are shared with pupils who complete their own self assessment. The reports are designed to be accessible, therefore they are short, precise and concise. They will also inform each pupil's transfer, transition and accreditation.

Accreditation comes through functional skills, entry level, AQA units of attainment, Asdan awards, Duke of Edinburgh awards, B.tech and GCSE. Pupils are entered for this accreditation and encouraged and supported to achieve at a level appropriate to them. There are no predicted

exam grades or league tables to present, but we still closely monitor each pupil's progress up and across the learning ladder. This is further supported by an enrichment grid map, which allows us to monitor their individual and collective opportunities in sport, the arts, community access and residential visits.

We have further data tracking records to monitor and support progress, but there are no level discussions with pupils. The records provide evidence for Ofsted and ensure learning pathways are open, rigorously monitored and accessible. Staff are still diligent with their data assessments: they refer not just to levels, but also the 'small steps' descriptors gained within and above the levels for each subject and student; end of key stage predictions are also made and regularly reviewed. But the data is never used as a label or measure.

As a special school we believe that achievements come through community cohesion, life skills, progression paths, student voice and collaborative working. We are very proud of our pupils' achievement, and so are they; they have a 'can do' attitude, recognising that achievement can be both individual and collective. It is built on a determination to learn despite often indescribable barriers. It is sustained by us all.

Reena Keeble DBE
Cannon Lane Primary School

Cannon Lane Primary is a large school in Harrow, London. It serves a diverse community, with two thirds of students coming from minority ethnic backgrounds. It was rated as 'outstanding' at its last Ofsted inspection and its head teacher is Dr Reena Keeble DBE. The school has recently amalgamated from separate junior and infant schools into an all through primary.

At Cannon Lane, we had regarded levels as fairly meaningless for some time. It had been obvious to us that a 2A in one school was not the same in another school, so we welcomed the opportunity to develop something that reflected our personalised learning approach. More importantly we wanted to find a way of reporting that made sense to our parents. Reporting a level did not tell

our parents what their child could or could not do, let alone their next steps in learning. I think this is very much our 'can do' approach. We relished the challenges but knew there would be barriers to have to overcome; mainly a mind shift for the staff of moving away from levels.

Our starting point for assessment was the National Curriculum. Using the first draft of the new curriculum in 2012, we broke down the expectations for pupils in English and maths into statements. Teachers planned to these statements and recorded achievement against each statement once it had been achieved. Teachers said this made their planning much easier and they knew exactly what each child had to do to achieve the statements. Therefore a very strong link between planning and assessment was established. Teachers found planning much easier and less time consuming and it also enabled us to strengthen our personalised approach to learning for our pupils. The teachers simply recorded a tick against each statement once it had been achieved. The beauty of this system was that it showed progression from Reception. For example, there were 13 statements in English for Reception. Year 1 statements began with point 14; so if a Reception child had achieved 16 statements in writing it was easy to see he/she was working to Year 1 expectations. The results were recorded on a simple spreadsheet and RAG (red, amber, green) rated. As a staff we developed our own expectations of where we expected children to be at particular time of the year. This was subjective and based on our own high expectations (which some said were too high).

Making the basic judgements of attainment was therefore easy because the statements we drew up were SMART; everyone had a shared understanding of them and what the evidence would look like. In addition, staff moderated and sampled work to ensure statements were being assessed in a consistent way. Data in the first year was recorded on a spreadsheet, but we now use a software package which we have helped to build with Classroom Monitor.

We can also use this approach to track specific groups of pupils. We do use numerical data in working at expected/exceeding/below levels.

However we have encountered two problems. Firstly, expectations for Years 3 and 4 and Years 5 and 6 are grouped together in the national curriculum, making it difficult to show progression from Year 3 to Year 4 or from Year 5 to Year 6. Secondly, our system was designed to show progress at the end of Year 4 and Year 6 in terms of the percentage that exceeded or were below or working at the expectations for that particular year. This is not robust enough and we need something much finer. We are currently working towards a solution and should have this sorted by the end of February/early March.

A key role for assessment is to keep parents informed. We are an 'achievement for all' school and use their system of structured conversations for all our parents each term. This means every parent gets to spend 20-25 minutes with their child's teacher (with the child being present if the parent chooses) to discuss progress and next steps. English and maths are discussed every term, as is the child's social and emotional development and dispositions and attitudes. Targets for the school and home to work on are agreed and a summary of the meeting is given to each parent. In the spring term, teachers also discuss progress in half of the foundation subjects and the remainder of the subjects are discussed in the summer term. In

addition to this parents are given a termly report on what their child has achieved and how they are developing in English and maths. These reports are generated by our software system. (We are currently developing a tool for parents to be able to access their child's report using a secure password rather than getting a hard copy.) We also send all the children's books home before the structured conversation so that parents have triangulated evidence of their child's progress – the books, the meeting and the written report. Parents no longer get a long school report at the end of the year and we have done away with parents' evenings because all structured conversations take place during the school day. Interestingly our parents prefer this and say that, as long as they have enough notice, they do not mind taking time off work.

Thinking ahead to implementation of the new approach to assessment, I think schools will probably use an assessment tool from possible case studies and make it their own. The challenge will come in ensuring the assessment procedures are robust enough. Implementing the tool is almost the easy part of assessment; it is the school procedures that need to be thought of. So, in conclusion, schools need to think not just about the specific tool they are using, but also their assessment processes, skills, systems and culture.

History and context of assessment

A brief history of national curriculum assessment and reporting

The current system of National Curriculum assessment was effectively the outcome of a debate set in motion in 1976 by the then prime minister James Callaghan in a speech at Ruskin College which created the notion of an entitlement to a National Curriculum and accountability for schools.

However, it was not until the Education Reform Act (ERA, 1988) that a National Curriculum and an associated assessment regime were introduced in England for the first time. Even in a brief history, it is worth mentioning the National Curriculum

Task Group on Assessment and Testing (TGAT) report as it made a significant contribution to the structure of the ERA and introduced the notion of tests and teacher assessment, National Curriculum levels and key stages, formative and summative assessment approaches and the use of tests to evaluate performance.

Kenneth Baker, the Secretary of State at the time, selected elements of the report favouring externally set tests and tasks in preference to the more formative function of assessment.

The National Curriculum Council (NCC) and the Schools Examination and Assessment Council

(SEAC)¹, both established under the 1998 Act, commissioned existing research and publishing organisations associated with universities to develop and trial a number of new assessment approaches. These were developed and phased in over a number of years, starting with Key Stage 1 then Key Stage 2 and finally Key Stage 3. With regards to Key Stage 4, GCSE syllabuses were developed to reflect the statutory requirements of the National Curriculum.

The initial tests and tasks were regarded by some as innovative and involved an element of practical work. All were marked by teachers. Some were better received than others, but ultimately all were met with a great deal of suspicion on the part of schools, particularly because of the workload. However, opposition to the tests had grown to such an extent that by the time Key Stage 3 went live in 1993 it led to a national boycott of the tests.

The initial level descriptions covered ten levels and were more complex than those used in 2013. Levels were presented under attainment targets each being assessed and weighted separately. As an example, the first National Curriculum for science had 17 attainment targets with 14 in mathematics. These attainment targets were split into statements of attainment and each statement of attainment was weighted separately. In addition to the weightings, some statements of attainment were necessary to achieve a level, others were not. The final level awarded to a pupil was therefore reliant on a complex procedure often resulting in a level that some teachers felt was not an adequate reflection of pupil performance.

1 The NCC and SEAC were abolished in 1993 and replaced by the School Curriculum and Assessment Authority (SCAA). In 1997, SCAA and the National Council for Vocational Qualifications (NCVQ) were abolished and replaced by the Qualifications and Curriculum Authority (QCA). In 2004, the QCA transferred all national curriculum assessment responsibilities to the National Assessment Agency (NAA). In 2008, the NAA was closed following test delivery failures and all national curriculum assessment responsibilities were brought back into QCA. In 2007 Ofqual was formed taking on the regulation of examinations and qualifications previously the responsibility of QCA. QCA was reconfigured and became the Qualifications Curriculum and Development Agency (QCDA). QCA was formally dissolved in 2010 when QCDA and Ofqual gained statutory status. In 2011, QCDA was officially closed and all national curriculum assessment functions were transferred to the Standards and Testing Agency, a part of the Department for Education.

A further contributing factor to the teachers' dissatisfaction with assessment arrangements was the government's decision to introduce performance tables. The first secondary school² performance tables were published by the department for education (DfE)³ in 1992 covering GCSE data and average point scores for 17-year olds entered for at least one GCE A/AS examination. The first primary school⁴ performance tables were published in 1996, showing the percentage of pupils eligible for Key Stage 2 assessment achieving level 4 in tests and teacher assessment in each of the three core subject, plus contextual data on pupil numbers and pupils with SEN. In 1997, publication of the primary tables was delegated to Local Education Authorities (LEAs) and by 1999, teacher assessment data had been dropped from the tables. However, teacher assessment data continued to be collected by the DfE but were not published until 2010 when new teacher assessment indicators showing the percentage of pupils on Level 4+ in English, maths and science were included.

These changes in data collection and publication illustrate the way the content of performance data has differed year-on-year.

Most of the changes have reflected policy decisions, but there have been other reasons for changing requirements. For example results were impacted by the outbreak of foot and mouth disease in 2001 and in 2002 post-16 data were delayed because of the Tomlinson enquiry in to GCSE A level results. In 2004, publication of Key Stage 3 English data was delayed until March 2005 because of a 'data collection disaster'. Problems re-occurred in 2005 when Key Stage 3 tables were again published late in March (2006) because of 'continued difficulties arising from late English marking and reviews'. Indeed, this was repeated in 2006 when tables were published in March 2007. In 2008 Key Stage 2 test data were delayed until April 2009 because of the Educational Testing

2 All maintained mainstream and special schools were included. Independent schools were included on a voluntary basis.

3 The current Department for Education has held differing titles over the period of this history. For ease, the term DfE has been used throughout.

4 Primary tables covered all maintained mainstream schools with KS2 pupils on roll (no special or independent schools).

Service's (ETS) failure in administration of the marking and as Key Stage 3 tests were to be withdrawn from 2009, the 2008 results were not published. In 2010, results for around 26 per cent schools that boycotted the tests were missing from the tables, the same year as science tests ceased to be used with all pupils and instead were used in a representative sample of schools to monitor national standards.

Other changes have had less obvious reasons. In 2002 Key Stage 3 results were published for all schools including independent schools. In 2003, Key Stage 3 results for independent schools were discontinued and the tables were once again published by the DfE. And in 2004 the tables were renamed as achievement and attainment tables.

In 2011, the numbers of pupils gaining 5+ GCSE A* to C grades including English and maths were published for secondary schools and contextual value added (CVA) was abolished with value added (VA) being re-introduced for Key Stage 1 and 2 mathematics and English.

More recent changes to the performance tables have resulted from the independent review of Key Stage 2 testing, assessment and accountability (Lord Bew, 2011): these are noted below. Full details of the achievement and attainment data and annual data specifications can now be found under the performance tables section on the DfE website.

A number of significant events and reports have also shaped the way National Curriculum assessments have changed and developed. Even from the earliest days of the National Curriculum, changes to the programmes of study and assessment requirements have been made. From the word go, the statutory orders that laid out the subject content of the National Curriculum proved contentious with all subjects attracting criticism from the teaching profession. With the exception of physical education, subjects were presented in a consistent manner organised through key stages, attainment targets, statements of attainment and programmes of study. The initial view was that the National Curriculum should occupy no more than 70-80 per cent of school time. However, the orders had been put together by subject experts in

isolation and without reference to other subjects. This resulted in duplication, over-subscription and complexity. This prompted ad-hoc reviews of subjects to reduce content; for example in 1993, the NCC had recommended revisions to English reducing the attainment targets from five to three and for design and technology from four to two.

In April 1993 John Patten, the then Secretary of State for Education, initiated a review of the National Curriculum and assessment framework. The grounds for the review were that the curriculum and assessment arrangements were not effectively integrated, that the curriculum had become over-extended and the assessment arrangements unduly complex. The NCC and SEAC were replaced by SCAA in October of that year under the chairmanship of Sir Ron Dearing who was charged to lead the review of a system he described as unjustifiably complex. Extensive consultations took place resulting in a reduction in curriculum content, fewer attainment targets and statements of attainment. Sir Ron recommended that in future, National Curriculum orders should be revised together rather than sequentially. Changes were also made to the tests and administration of the assessments. Tests were only maintained in English, mathematics and science with teacher assessments, at least in theory, given equal weighting. The ten level scale was reduced to eight.

In the same year, it was announced that Key Stage 3 tests would be externally marked. In the following year, external marking was introduced across key stages two and three and national data collection was introduced. In 1996 the DfE published the first primary performance tables as noted above.

The following few years saw further consultation and year on year minor changes: the boycotting of tests faded away.

In 1997 the newly elected Labour government introduced the national literacy strategy and the national numeracy strategy the following year. To the Strategies the tests were a data source to measure their progress on which they established a system of pupil, teacher, school, local authority and national targets. To the government, the

targets were seen as a device to drive up and measure standards. David Blunkett, the then Secretary of State for Education staked his position on the impact of National Strategies by promising to resign if the target of 80 per cent of pupils achieving a level 4 or higher in English by 2002 was not met. Given that in 1997 the number achieving this level was 63 percent, this brought even more pressure on what was already a high stakes environment.

1998 saw the first of three major national failures in the marking and publication of results system.

This failure attracted allegations in the Daily Telegraph that the Secretary of State had manipulated the level boundaries to achieve higher performance against his targets.

This led to the Rose Report *Weighing the Baby* in 1999. David Blunkett and QCA were exonerated together with a recommended moratorium on changes to the tests and a review of the purposes of assessment. In common with later failures, concerns were raised about inadequate project management, not acting on concerns expressed by staff, untrained staff and poor communication between QCA and the department for education. Of interest, the major delivery failures have usually followed the appointment of new contractors and/or changes to the system to meet ever increasing data requirements. Despite this, the DfE was reluctant to introduce on-screen marking which was seen by QCA as a means to improve marking quality and shorten processing time.

In 2003, the National Assessment Agency was formed as a means of separating the delivery and regulatory functions of QCA. Further changes to the tests and mark schemes following the Rose Report and the subsequent review of assessment were also implemented.

Evidence was beginning to emerge that the government had not achieved its targets for primary schools as set by the strategies. In 2003 the Secretary of State, Charles Clarke, announced that in response to concerns raised by the profession about excessive pressure caused by top down targets, the primary targets would go and a pilot of tests and tasks at Key Stage 1 informing

teacher assessment would be run in 2004. In 2005, Key Stage 1 moved to teacher assessment as the prime indicator of achievement. Further, the national target of 85 per cent of 11-year old pupils achieving at least level 4 was relaxed to be met 'as soon as possible'. Interestingly in Wales, Key Stage 1 tests were abolished in 2001.

In 2004, the NAA introduced component marking for Key Stage 3 English where reading and writing were assessed by different markers as a means of reducing the burden on markers and improving marking quality. The reading and writing test scripts were then re-united for borderlining, a quality check for scripts adjacent to performance level boundaries. However, the new system slowed down the marking process and resulted in delays to the delivery of results to schools. Mike Beasley, a QCA board member, undertook a review of the delivery failure which repeated some of the findings of the 1998 failures. However, whilst the report found no reason to doubt the test, quality of marking or the final results, it criticised the delivery system as being badly flawed characterised by poor leadership and inadequate project management. Changes were made to the management systems for 2005, but concerns over marking quality had heightened. To allay fears, the NAA established marking centres, but this in turn delayed Key Stage 3 results again in 2005 and 2006.

Faced with continuing problems with its targets, the government launched the *Making Good Progress* consultation in 2006. The consultation explored a system whereby schools could focus more systematically on assessing the progress of pupils supported by a system of more flexibly timed statutory tests. In 2007 the NAA was commissioned to develop and pilot single level tests: one-level, summative, reportable measures of pupil performance that could be administered when teachers considered pupils were ready to take the test. In the same year the Chief Executive of QCA, Ken Boston criticised practice testing in schools, saying "*The key to driving up performance at key stages 2 and 3 is better teaching based on diagnostic assessment and personalised learning, not more practice drill in taking tests.*" However, the single level test system was never implemented.

Around the same period, the QCA was developing the Assessing Pupils' Progress (APP) initiative with its focus on supporting teacher assessment. APP materials were available to schools along with optional tests provided by QC(D)A which were designed to help teachers identify pupils' strengths and weaknesses and guidance materials such as the Understanding Progress series. These materials provided a greater focus on assessment as a tool to support teaching and learning. However, tests continued to dominate the agenda.

In 2008, the test system saw its most dramatic delivery failure. The contract for the delivery and marking of tests had been awarded to ETS a USA based assessment company a year earlier giving them one year to prepare their systems. In the light of ETS' failure to deliver test results to schedule, the Qualifications and Examinations Regulator (Ofqual) and Ed Balls, the then Secretary of State for Education, remitted Lord Sutherland of Houndwood to lead an independent inquiry. By the time the inquiry reported in late 2008, Ed Balls had announced the end of Key Stage 3 tests for 2009 and single level tests. Lord Sutherland's report recommended that test delivery should be modernised, with any new processes piloted and that customer service to schools should be vastly improved. In particular, the report was highly critical of the due diligence process during the procurement exercise and the ambiguous role of the NAA within QCA's corporate structure. The report also called for more clarity of Ofqual's role and a strengthening of its resources and skills to monitor QCA more thoroughly. The recommendations of the inquiry were accepted by the Secretary of State and Ofqual and fully implemented by QCA. The NAA was discontinued and the management of the tests was located fully within QCA. The Chief Executive of QCA resigned. From 2009, the test delivery and marking process has seen considerable change with year-on-year improvements to the service.

Nevertheless, concerns over the more fundamental aspects of testing continued. In 2008, as part of the announcement to discontinue testing at Key Stage 3, the Secretary of State established the expert group on assessment with a remit covering Key Stages 1, 2 and 3 and a focus on the

essential purposes of assessment and how they could be best met. The Group reported in 2009 with a range of recommendations including cross key stage moderation, strengthening the quality of teacher assessment and national sample tests at Key Stage 3 to monitor standards over time. However, other than the development of sample tests, the recommendations were overtaken by the arrival of a new coalition government in 2010.

It was not long before Michael Gove, Secretary of State for Education, announced the *Review into Key Stage 2 testing, assessment and accountability* chaired by Lord Paul Bew. The review reported in July 2011 with a range of recommendations. The report supported external school-level accountability and a focus on progression but recommended greater emphasis on teacher assessment within statutory assessment. With regards to the latter, the review recommended that English reading tests should continue to be externally marked but that writing composition should be subject only to summative teacher assessment with spelling, punctuation grammar and vocabulary assessed through an externally marked test. Cluster moderation by schools was encouraged as was the wider use of Key Stage 2 pupil-level data by secondary schools, but this was not granted the weight of a recommendation. Mathematics as an externally marked test and the continued use of national sampling of science were recommended. The report was accepted by the government and affectively underpins the current National Curriculum assessment system for Key Stage 2.

The Secretary of State also announced the closure of QCDA and a review of the National Curriculum by an expert panel. The National Curriculum Review Expert Panel reported in 2011 and as a part of their deliberations they expressed a view that programmes of study and attainment targets have often lacked precision even after successive reviews of the content of the National Curriculum. This led the Panel to conclude that attainment targets and level descriptions should not be retained in the revised National Curriculum. In June 2013, the Secretary of State announced that levels had become too abstract, detracted from real feedback to pupils and parents and that schools have found difficulty in applying them consistently.

So from 2014, levels would be removed from National Curriculum assessments. New tests based on the new National Curriculum will be taken by pupils for the first time in the summer of 2016.

The announcement was accompanied by the launch of a consultation on *Primary assessment and accountability under the new National Curriculum*. The consultation proposed amongst other things:

more demanding tests, a baseline test at the end of Key Stage 1 or at the start of reception, test results reported using a scaled score, comparison of pupils against the national cohort by decile and threshold attainment measures at a much higher level.

The outcome of the consultation is expected in February 2014.

A brief history of GCE and GCSE

The General Certificate of Secondary Education (GCSE) was introduced for first teaching in 1986 and first examination in 1988. The GCSE was a single system of examinations designed to replace the dual system of the General Certificate of Education (GCE) Ordinary level and the Certificate of Secondary Education (CSE). The GCSE was available through six examination boards in England, Northern Ireland and Wales.

The GCE was established in 1951 and was available at ordinary, advanced and scholarship levels. Before 1951, the national school examination system offered the School and Higher School Certificate under the control of the Board of Education. Both certificates recognised attainment in defined groups of subjects with the School Certificate aimed at 16-year olds and Higher Certificate for 18-year olds. Entries for the examinations were small compared with current day standards. As early as 1943 the Norwood report had proposed a new system of single subject awards with an emphasis on teacher rather than exam board assessment particularly at the School Certificate level. This proposal was not accepted but following the 1944 Education Act which raised the school leaving age to 15 and established a system of grammar, technical and secondary modern schools, the GCE was introduced with its single subject offer. However, the GCE was exclusively aimed at the top 20 per cent of the population and as such became a predominantly grammar school qualification. Indeed technical and secondary modern schools were discouraged from large scale GCE entries which resulted in the creation of a proliferation of localised leaving certificates resulting in the majority of

pupils leaving school with no nationally recognised qualification. As for GCE A levels, they were even more exclusive aimed at a proportion of those taking the ordinary level. As such it became known as the gold standard and has since its inception acted as the key entry qualification for universities.

It was not until the 1960s that the CSE came into being. The Schools Council was established in 1964 to replace the Secondary Schools Examinations Council with a remit to oversee the developing examination system. The CSE was introduced in 1965 and was available in three formats: Mode 1 where syllabuses and examinations were set and marked by an examination board; Mode 2 where schools set their own syllabus but examinations were set and marked by a board; and Mode 3 where schools set the syllabus and examination and marked their own examinations subject to approval and moderation by a board. The CSE was delivered by fourteen regional examination bodies and were targeted at the 40 per cent of pupils below the 20 per cent GCE ordinary level target. These fourteen bodies were in addition to the eight GCE boards.

The GCE ordinary level was graded A to E with grade A being the highest. The CSE was graded on a five point scale, 1 to 5, with grade 1 being the highest and regarded as equivalent to a grade A to C at the GCE ordinary level. Unlike most GCEs, the CSE had an emphasis on the assessment of work undertaken during the course of study, rather than through a one-off final examination. However, with the school leaving age being raised to 16, more children were entered for

the two examinations and questions of comparability grew as did the confusion about the two grading systems by employers and the general public. As early as 1971, the Schools Council were looking at the possibility of combining the two qualifications which resulted in a joint 16+ pilot courses up to 1974 which were jointly developed by GCE and CSE boards. Most of these were discontinued at that point except for those offered by the Northern Examining Association which ran up to 1987.

Following a number of feasibility studies, debates and changes in government, the GCSE was approved in 1984. Subject specific and general criteria were developed which the new syllabuses were required to meet and in 1986, the first teaching began with the first examinations being held in the summer of 1988. The criteria represented agreed standards rather than a reliance on norm referencing, or pre-set quotas, which was a feature of qualifications up to that point. Of equal significance, the 1988 Education Reform Act introduced the Secondary Examinations and Assessment Council which was granted statutory powers over assessment arrangements. This was effectively the start of a more centrally controlled and ultimately regulated examination system.

Unlike the GCE ordinary level and the CSE, the GCSE was not specified for a particular ability group with the intention that around 90% of the school population should achieve the standard of 'average' which at the time was grade F. Because of the wide ability range expected to take the GCSE, examination papers were designed in a new way. A key aim of the GCSE was that of allowing candidates to demonstrate what they know, understand and can do. In order to achieve this aim, differentiated, or 'tiered' questions papers were used.

Assessed coursework was a feature of the GCSE with the initial syllabuses having at least 20 per cent with some subjects going well above the minimum. English syllabuses for example contained between 50 to 100 per cent coursework. However, this came under attack with the then prime minister John Major calling for 20 per cent as the upper limit in 1992. As a result coursework was settled between 20 and 40 per cent. This was challenged by Sir Ron Dearing in the mid-1990s who recommended

an increase, but this was not accepted by Gillian Shepherd, the then Secretary of State for Education. Along with comparability of standards over time, the place of coursework has attracted considerable debate throughout the lifespan of the GCSE. In 2006, the QCA raised concern over the fairness of coursework citing problems such as parental support, cheating and plagiarism – particularly referring to the internet as source. To counter this, ministers announced in 2009 that controlled assessments would replace the more open approach of coursework.

Following a consultation between June and September 2013, Ofqual announced in November that further modifications to the GCSE would be introduced impacting qualifications for first teaching from September 2015. The key features include a new 1-9 grading scale (with 9 being top); limitations on the use of tiered papers; an end to modular examinations – all GCSEs will be examined at the end of the course; examinations will be the 'default' method of assessment unless there are issues of validity; and all examinations will be held in the summer with the exception of English language and mathematics, where there will also be exams in November for students who were at least 16 on the preceding 31 August.

At the time of writing, further announcements are expected from Ofqual covering decisions on non-exam assessment on a subject-by-subject basis and whether November examinations should be made available for 16-year-olds on the preceding 31 August in a wider range of subjects.

It was also confirmed that from 2015, English language will be un-tiered and fully assessed by an external examination. As now, speaking assessment will be reported separately. English literature will also be un-tiered and assessed by an external examination. For mathematics, overlapping tiered papers will be used and assessed by an external examination.

Assessment practice: an international snapshot

One of the recurring themes in the evidence presented to the Commission has been that of trust in teachers' assessments. There are two main reasons underpinning the concerns expressed:

1. the unintended consequences of a high stakes assessment system, for example the pressures exerted by the publication of performance tables which have created a perverse incentive to inflate assessments; and
2. concern over the level of assessment expertise in the teaching profession.

As a result, teacher assessment has been downplayed in the accountability system in England. The following sections provide a snapshot of assessment practice in other countries.

The sections on **Sweden, Australia, Norway, and New Zealand** are taken from the *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes*.

The **Swedish** approach combines national standard-setting and central test development with a high degree of trust in school professionals to carry out evaluation and assessment.

Many evaluation and assessment activities including student assessment, teacher appraisal and school quality reporting are managed internally at the school level. This approach fosters and encourages school leader and teacher professionalism in evaluation and assessment.

While a lot of quality assurance work happens locally and informally, these practices are frequently not documented and there is little evidence as to whether good practice is spread and shared across the system.

There is a strong focus on classroom-based assessments through which teachers collect a variety of evidence on student progress and provide regular feedback to students. National tests at key stages of education are intended to capture a wide range of curriculum goals through performance-based tasks including oral assessment and team projects. The tests are summative in Year 9 and upper secondary school and intend to provide a more standardised and external measure of student achievement.

However, as all other types of assessment in Sweden, the national tests are corrected and graded by the students' own teachers, and the weight of test results in students' grades is determined locally.

This raises concerns about inequities in grading. In fact, teachers' marking of the performance-based national tests has shown to be uneven. Possible explanations are that grading criteria are not adequately detailed and that teachers vary in their capacity to score student achievement on performance-based tests. There is a lack of external reference points and moderation to ensure that student assessment in Sweden is reliable and fair.

Source: OECD REVIEWS OF EVALUATION AND ASSESSMENT IN EDUCATION: SWEDEN
© OECD 2011

The **Australian** approach combines the development of goals, monitoring and reporting at the national level with local evaluation and assessment practices shaped by jurisdiction-level school improvement frameworks.

The current strategy for student assessment consists of a combination of National Assessment Program Literacy and Numeracy (NAPLAN) and teacher-based assessments against the full range of curriculum goals. The latter implies a considerable investment on teacher capacity to assess against the standards, including specific training for teachers, the development of grading criteria and the strengthening of moderation processes within and across schools. Also, the current prominence of NAPLAN within the student assessment framework requires particular care about not reducing the importance of teacher-based assessment.

Teachers benefit from a high degree of trust and extensive autonomy, but they have few opportunities for professional feedback.

Source: OECD REVIEWS OF EVALUATION AND ASSESSMENT IN EDUCATION: AUSTRALIA
© OECD 2011

Norway has a well-established tradition of decentralisation and school autonomy, with a strong sense of individual schools being 'owned' by their local communities and accountable to them rather than the national authorities.

The Norwegian authorities have set up a national quality assessment system (NKVS) for the education sector in 2004. NKVS provides access to a range of data and tools intended to help schools, school owners and education authorities evaluate their performance and inform strategies for improvement. The system initially included mandatory national student assessments, user surveys and a web-based school portal, and was later complemented by additional tools and guidance to support evaluation at the local level.

The successful implementation of an evaluation and assessment framework crucially depends on whether professionals in counties, municipalities and schools have the understanding and competencies to collect, analyse and interpret evaluative information with a view to improve practices. Embedding an evaluation culture in schools and municipalities across Norway is a large culture shift that requires further investment in professional learning opportunities, targeted to the needs of different stakeholder groups.

Source: OECD REVIEWS OF EVALUATION AND ASSESSMENT IN EDUCATION: NORWAY
© OECD 2011

New Zealand has developed its own distinctive model of evaluation and assessment characterised by a high level of trust in schools and school professionals. There are no full-cohort national tests and teachers are given prime responsibility to assess their students' learning.

National Standards were introduced in primary education in 2010 to provide clear expectations for student learning in mathematics, reading and writing and help teachers make and report overall teacher judgements (OTJs) based on a range of assessment evidence. In a context where there is a general consensus against national testing in primary education, the introduction of Standards is seen as an alternative way to make information about student learning more consistent and comparable. However, further developments are necessary to embed the Standards within the primary school system.

These include (1) Ongoing investment in teacher professional development to build teachers' capacity to assess students in relation to the National Standards; (2) Stronger support for

systematic moderation processes to ensure that OTJs are reliable and nationally consistent; (3) Better articulation between the National Standards, the National Curriculum and existing assessment tools; (4) Clearer statements regarding the kind of information that standards-based reporting can and cannot provide and the uses of reporting information that are considered appropriate; and (5) Further work to ensure that the Standards' focus on literacy and numeracy does not marginalise other learning areas where measurement of performance and progress is more challenging.

In the context of self-management, individual schools can be relatively isolated and have limited opportunities for collegial networking and peer learning. There are a range of policy options to strengthen the connectedness of schools and help spread and share effective evaluation and assessment practice. These include (1) Providing cluster funding for groups of schools to pool evaluative information and engage in collaborative analysis and interpretation of data; (2) Supporting the collaboration of schools with an external facilitator or 'critical friend' such as a professional development provider; (3) Relying as much as possible on practitioners in the role of peer evaluators or participating in ERO review teams; and (4) Building further on recent developments to strengthen the Regional Offices of the Ministry of Education and enhancing regionally based school support structures.

While there has been strong focus on building evaluation and assessment competencies at the school level, further investment in professional development is necessary to ensure that practices are consistently effective across New Zealand. Teachers need to develop not only the capacity to use, interpret and follow up on results obtained from nationally provided assessment tools, but also to develop their own valid and reliable assessment tools, adapt assessment to diverse learner profiles and communicate and report assessment results effectively.

Alongside general training in assessment literacy, teachers and school leaders also need to further develop skills to collect school-wide assessment data; disaggregate data for relevant sub-groups; and interpret and translate assessment information into improvement strategies. Central agencies

could consider developing a unique set of teacher competencies in assessment to set clear targets for initial teacher education and continuing professional learning. Given the key role of school leaders in New Zealand's devolved education system, there is also a need to firmly embed a focus on effective evaluation and assessment in the competency description, training, performance appraisal and support materials for school leaders.

Source: OECD REVIEWS OF EVALUATION AND ASSESSMENT IN EDUCATION: NEW ZEALAND © OECD 2012

In **Finland** school inspections were abolished in the early 1990s. The ideology is to steer through information, support and funding. The activities of education providers are guided by objectives laid down in legislation as well because the national core curricula and qualification requirements. The system relies on the proficiency of teachers and other personnel.

There is strong focus on both self-evaluation of schools and education providers and national evaluations of learning outcomes. National evaluations of learning outcomes are done regularly, so that there is a test every year either in mother tongue and literature or mathematics. Other subjects are evaluated according to the evaluation plan of the Ministry of Education and Culture. Not only academic subjects are evaluated but also subjects such as arts and crafts and cross-curricular themes.

From the schools' perspective, the evaluations are not regular as they are sample-based. The education providers receive their own results to be used for development purposes.

The main aim of the national evaluations of learning outcomes is to follow at national level how well the objectives have been reached as set in the core curricula and qualification requirements. Consequently, the results are not used for ranking the schools.

The main type of pupil assessment is the continuous assessment during the course of studies and final assessment. Continuous assessment is to guide and help pupils in their learning process. Each student receives a report at least once every school year.

There are no national tests for pupils in basic education in Finland. Instead, teachers are responsible for the assessment in their respective subjects on the basis of the objectives written into the curriculum. Also the grades in the basic education certificate, the final certificate given at the end of year 9, are given by the teachers.

On the basis of this assessment pupils will be selected for further studies. Therefore, the national core curriculum contains assessment guidelines in all common subjects. One task of basic education is to develop the pupils' capabilities for self-assessment. The purpose of this is to support the growth of self-knowledge and study skills and to help the pupils to learn to be aware of their progress and learning process.

The first national examination is at the end of general upper secondary education. General upper secondary education ends with a national matriculation examination, which comprises four compulsory tests: mother tongue and, according to each candidate's choice, three of the following: the second national language, a foreign language, mathematics or one subject in general studies, such as humanities and natural sciences. Students may also include optional tests.

Having completed the matriculation examination and the entire upper secondary school syllabus, students are awarded a separate certificate that shows details of the examinations passed and the levels and grades achieved.

Teachers in basic and general upper secondary education are required to hold a Master's degree. At most levels of education the teachers are required to participate in in-service training every year as part of their agreement on salaries.

Finnish teachers consider in-service training as a privilege and therefore participate actively.

The State also provides in-service training programmes, primarily in areas important for implementing education policy and reforms. The education providers can also apply for funding to improve the professional competence of their teaching personnel. Teachers are recognised as keys to quality

in education. Therefore continuous attention is paid to both their pre-service and continuing education.

Source: *Finnish education in a nutshell. 2012. Ministry of Education and Culture, Finnish National Board of Education, Centre for International Mobility (CIMO)*

The examples from **Singapore, Korea** and **Japan** that follow are drawn from the *International Review of Curriculum and Assessment Frameworks* (INCA) country archives which were set up by the Qualifications and Curriculum Authority and are now contained in the national archive. They may not reflect the latest practice in each country.

In **Singapore**, the assessment and qualification arrangements are fairly similar to those in England. At the end of primary education there is the national Primary School Leaving Examination (PSLE) whose primary purpose is to stream students by ability onto different courses in lower secondary (it also informs assessment of school performance). Continuous formative assessment in primary and lower secondary education is used by teachers to assess progress in extracurricular as well as academic studies. There are also mid and end of year examinations, which provide more summative information. A review in April 2009 is leading to less emphasis on these in lower primary, in favour of 'bite-sized' assessment to improve student confidence and engagement⁵.

All students wishing to go on to one of Singapore's two universities must also achieve the Scholastic Achievement Test (SAT) covering verbal reasoning and mathematics. The Centre for Testing and Assessment Pte. Ltd (CTA) has been set up by the universities to administer the tests. The Ministry of Education is responsible for the administration of all national tests⁶.

There are various types of assessment systems in **Korea**. These include the nationwide system of scholastic achievement tests (*SATs*) (the national assessment of educational achievement), and continuous classroom assessment by teachers. A revised system of periodic national assessments of student achievement (the national assessment of educational achievement) began to be implemented in September 2000, the principal aim of which is to

monitor the curriculum. Under the system, Korean language, mathematics, science and social studies are assessed every two years (two subjects each year), while English communications skills and the use of information technology skills are assessed once every three years. Small samples of students (between 0.5 per cent and one per cent of the whole student population in specific years/grades) are involved in the tests.⁷

Entry into higher education in Korea is based on students' high school records, extra-mural activities and scores in national tests (College Scholastic Achievement Tests – CSATs) as well as assessment arrangements by individual universities. These take place on one day a year (special arrangements are made to reduce distraction to candidates by staggering rush hour traffic and grounding flights)⁸.

Japan, like Sweden, has no formal system of assessment for pre-school education⁹. In primary and lower secondary there have been national tests, which are survey-like in style, happening last in 2007 and involving many but not all schools. This was to measure the national standard of education after structural reforms. Other assessment is teacher-driven which includes teacher devised formative assessment which may be criterion or norm referenced; involving a comparison of the performance of individuals with that of their peers and reported to parents in terms of grades. There are no standard procedures laid down nationally as to how such grades should be derived or described¹⁰. Students are also encouraged to assess their own and their peers' work¹¹. Completion of upper-secondary education is certified by individual principals with no external moderation; although there is a separate national university admission examination and universities may also administer their own admissions assessment¹².

5 INCA Singapore country archive. Section 6.2 (Available at: <http://www.inca.org.uk/1082.html>) and Section 6.3 (Available at: <http://www.inca.org.uk/1083.html>)

6 Ibid

7 INCA Korea country archive sections 6.4.2 and 6.4.4 (Available at: <http://www.inca.org.uk/1403.html>)

8 Ibid

9 INCA Japan country archive. Section 6.1 (Available at: <http://www.inca.org.uk/1477.html>)

10 INCA Japan country archive. Section 6.3 (Available at: <http://www.inca.org.uk/1481.html>)

11 INCA Japan country archive. Section 6.2 (Available at: <http://www.inca.org.uk/1478.html>) and Section 6.3 (Available at: <http://www.inca.org.uk/1481.html>)

12 INCA Japan country archive. Section 6.4 (Available at: <http://www.inca.org.uk/1483.html>)



1 Heath Square
Boltro Road
Haywards Heath
West Sussex
RH16 1BL

t: **0300 30 30 333**

w: **www.naht.org.uk**

© February 2014

Design and print: www.graphicimpressions.co.uk