

# Education Select Committee Primary Assessment Inquiry: NFER Response, 28<sup>th</sup> October 2016

Contact: Catherine Kirkup, [c.kirkup@nfer.ac.uk](mailto:c.kirkup@nfer.ac.uk)

Annex 1 contains more information about NFER and why we are submitting this response.

## Overview of our response

1. **We welcome the Select Committee's decision to hold an inquiry into primary assessment.** NFER has been developing high-quality, robust, reliable and informative assessments for 70 years, from individual classroom assessments to large-scale national assessments that support system-wide improvement. We have an extensive track record in undertaking research on assessment related issues (see, for example, Tisi et al, 2013). We believe that the use of reliable assessments is an important part of effective teaching and learning.
2. **Our key recommendations are:**
  - a) An assessment system needs to meet a number of different purposes for a variety of different stakeholders. If the principal purpose of primary assessment is one of school accountability, in the opinion of NFER, statutory national testing is the best assessment model to use as the basis for judging the performance of schools. Alongside this, teacher assessments continue to play an important, complementary, role as part of the wider assessment system.
  - b) We support the intention to measure progress from school entry and believe that it is possible to develop an age-appropriate, robust, standardised assessment for this purpose.
  - c) Over a relatively short period of time, there have been a large number of changes to the statutory national assessment system. We recommend that a robust evaluation is undertaken to assess the impact of the assessment reforms on teaching and learning, and to develop an agreed approach to implementing any changes in future that takes account of the needs and planning timescales of schools.
  - d) NFER believes that a shared understanding of assessment is inextricably linked with teachers' professional development. Consideration should be given to providing funding for a programme of support and training for assessment data literacy.

## The purpose of primary assessment and how well the current system meets this (including the advantages and disadvantages of assessing pupils at primary school)

3. **An assessment system needs to meet a number of different purposes for a variety of different stakeholders.** An effective system will therefore be made up of a combination of different forms or models of assessment, each serving one or more purposes. At both the school and system level, there should be a clear, and agreed, understanding of how the assessments will be used in order to achieve the desired outcomes.

4. **It is possible to identify a wide variety of different purposes for assessment:** Newton (2007) lists ‘a selection’ of eighteen. The following list highlights some that are relevant to current debates about assessment:
- to enable benchmarking between schools and as the basis for judging the performance of the school
  - to monitor standards over time (locally and nationally)
  - to summarise attainment at the end of a particular phase of education
  - as a baseline against which to measure future progress
  - to evaluate an individual pupil’s knowledge, skills and understanding that is helpful to the pupil, parent/carer and teachers in determining the teaching and learning needs of the pupil (formative assessment)
  - to diagnose particular learning difficulties (e.g. the Year 1 phonics screening check aims to identify pupils who need additional support in developing their phonic decoding skills)
  - to make comparisons about the effectiveness of the school systems of whole countries e.g. the OECD’s PISA (focused on the secondary phase), and the IEA’s TIMSS and PIRLS programmes.
5. **There is a range of different kinds of assessment that are more or less appropriate for achieving each of these purposes,** including summative (end of year or stage) tests, formative tests, teacher assessment, and sample tests. A balance needs to be struck between:
- proliferating different assessments that are geared to the specific requirements of each purpose and that then become a burden on schools and pupils
  - seeking to draw conclusions from assessments designed for one purpose in other contexts.
6. **It is unrealistic and inappropriate to expect assessments designed for one purpose to be as effective for others.** The focus of this inquiry is on the statutory assessment tests, which **are designed to be a part of the school accountability system.** Alongside the issue of how well the current National Curriculum Tests fulfil their purpose in terms of school accountability, there are questions about what other assessments are needed in primary schools, for example, to meet the needs of teaching and learning.
7. **The main forms of assessment currently in use within the primary phase are as follows:**
- **Statutory national tests,** at the end of KS1 and KS2, which provide reliable results about pupil attainment, allow local and national benchmarking and enable the government to hold schools to account for their performance. These assessments are carefully designed and rigorously trialled to ensure reliability and validity.
  - **Teacher assessment,** which provides specific information about pupil performance to inform teaching and learning throughout the year, to facilitate conversations with pupils and parents, and to complement test results.
  - **Externally provided tests** which, if carefully designed, can help teachers to benchmark their pupils’ performance against a national standard and to monitor progress. The outcomes of such tests may also contribute to teachers’ formative

assessment practices during the school year, by helping teachers identify the necessary next steps for pupil learning.

8. **The issue of teacher assessment complementing test results is an important one because tests (statutory or non-statutory) can only sample from the underlying curriculum;** it is not possible to test all aspects of the curriculum in one test and some aspects need a different form of assessment. For example, speaking and listening cannot be assessed in a written test of English. It is also an issue in Key Stage 2 science. The Test Framework published by the Standards and Testing Agency (STA) acknowledges that not all aspects of the ‘working scientifically’ element of the science curriculum can be assessed in the Key Stage 2 science sample test. See the test framework for further details:  
[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/439614/2016\\_KS2\\_Sciencesampling\\_framework\\_PDFA.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/439614/2016_KS2_Sciencesampling_framework_PDFA.pdf)
9. **Additionally, alongside cognitive aspects of pupil’s development, there are other characteristics of effective learning, such as resilience, engagement and motivation, which are more difficult to measure through a formal test.** Some schools may contribute to the development of these attributes within their pupils more effectively than others (see <https://www.gov.uk/government/news/funding-boost-for-schools-helping-pupils-develop-character>)
10. **Perceptions of the various forms of assessment may be coloured by the uses to which they are put and the extent to which they are ‘high-stakes’, either for the pupils or for the schools.** Concerns are sometimes raised about the use of National Curriculum Tests as a measure of school performance, particularly when decisions about a school, such as whether it is necessary for external agencies to intervene, will be made based on the performance of its pupils in these tests. However, very similar tests are widely used by schools on a voluntary basis when the stakes are not as high. This is because they provide a reliable way for teachers to track how pupils are performing (and don’t require schools to spend time developing their own tests<sup>1</sup>). Some of the concerns expressed about the nature of external assessments reflect these issues rather than the nature of the assessments themselves.
11. **Alternative approaches to primary assessment could be considered but would not provide the same level of accurate or reliable data that enables the government to hold schools to account.** Some countries do not use a national standardised assessment system during the primary phase of education. For example, in Finland,

---

<sup>1</sup> The Government’s recent teacher workload challenge Data Management Review Group (2016) identified that the use of rigorous externally developed assessment could contribute to reducing teacher workloads. Their report states:

Standardised tasks and test items can be a useful tool in supplementing teacher judgement and assessment. .... Used well, as part of an overall assessment scheme (particularly when accessing resources of an established quality) they can offer a quicker and less burdensome way for a teacher to check progress, to feed into a wider professional conversation.

during compulsory primary education, ‘teachers set their own assessments based on national core curriculum descriptors of ‘good’ performance’ (A rapid evidence assessment on the impact of curriculum and assessment arrangements within high performing countries, Welsh Government Social Research, 2013: p.37) However, there is evidence that the removal of school performance measures can reduce school effectiveness (Burgess, S, Wilson, D, and Worth, J., 2013, although note that this is focused on the secondary phase). Additionally, although teacher assessment plays a central role in any assessment system, there are concerns that teacher assessment can be biased [see <http://www.cls.ioe.ac.uk/shared/get-file.ashx?itemtype=document&id=1715%20>]. By contrast standardised assessments such as the National Curriculum Tests at Key Stage 1 and 2 have the advantage of providing objective and robust external measures of pupil performance which are not subject to the same potential biases as teacher assessment.

12. **If monitoring national standards was the key driver of primary assessment, national sample tests (such as those used to monitor KS2 standards in science) could be used to monitor standards over time.** However, a sampling model would not provide information about the performance of all pupils or all schools and therefore would be inappropriate for school accountability purposes.
13. **If the principal purpose of primary assessment is one of school accountability, in the opinion of NFER, statutory national testing is the best assessment model to use as the basis for judging the performance of schools.**
14. If national testing is to be used for accountability, the supplementary questions to be considered are: which areas of the curriculum should be assessed; at what ages should the assessments take place; and how should the results be reported in order to maximise their usefulness to different stakeholders within the system:

- **Which areas of the curriculum should be assessed?**

The current focus on English and mathematics is based on the premise that attainment in these subjects underpins access to all other areas of the curriculum, and in turn leads to better education and employment opportunities in the future.

However, assessment used for accountability can distort the priorities of the school. There is some evidence that assessing only a few subjects can lead to a narrowing of the curriculum as teachers focus on these subjects at the detriment of other subjects that are not assessed through a high-stakes test. This issue was identified by Lord Bew in his review into primary school testing and accountability in 2011. <http://www.educationengland.org.uk/documents/pdfs/2011-bew-report-ks2tests.pdf>

As we have argued previously (NFER’s response to the Primary Assessment and Accountability under the New National Curriculum, 2013) the emphasis on English and mathematics is important, but equally it is vital to recognise that primary schools have a much broader focus than this. Other aspects could also be reported on to show they are valued in schools. In 2009, NFER published a paper on the information published about schools in other countries and made recommendations

about what ought to be published about schools in England (see Maughan *et al.*, 2009).

Teacher assessment can be used to complement the tests in English and mathematics, to assess other areas of the curriculum and to provide a broader picture of pupil attainment. National testing in further subjects would increase the costs of development of the key stage tests and increase the burden of testing for pupils and schools.

- **At what ages should the assessments take place?**

As part of the reforms to accountability, the new primary floor standards can be met by schools either in terms of their pupil attainment or the progress their pupils have made. A drawback of the current system is that progress is measured only from the end of KS1 to the end of KS2, ignoring the valuable input schools can make during the first three years of primary education. **We support the intention to measure progress from school entry and believe that it is possible to develop an age-appropriate, robust, standardised assessment for this purpose.**

- **How should the results be reported in order to maximise their usefulness to different stakeholders within the system?**

In NFER's view, measurement should be used to drive an improvement in standards and therefore it is vital to have effective feedback at every level in the system. At the school level, such feedback needs to be in a form that teachers can use and that will be understood by parents / carers and pupils.

Following the removal of national curriculum "levels", changes have been made to the way in which the key stage test results are reported. The results are now reported as scaled scores (see <https://www.gov.uk/guidance/scaled-scores-at-key-stage-2>). There are also performance descriptors to inform statutory teacher assessment judgements at the end of Key Stages 1 and 2.

The decision about the abolition of levels included consideration of how well parents / carers understood the system. It is perhaps too early to say whether the introduction of scaled scores has affected understanding of the key stage results (by both teachers and parents/ carers) but **once the new system of reporting has had some time to bed in, consideration should be given to an evaluation of the new reporting mechanisms.** Such an evaluation could also investigate the training and support needs of teachers and school leaders (see also below).

## **How the most recent reforms have affected teaching and learning**

15. **Over a relatively short period of time, there have been a large number of changes to the statutory national assessment system** including: the statutory phonics screening check in year 1; the introduction of a new test (grammar, punctuation and spelling); more challenging test content / standards (linked to the revised 2014 national curriculum); and new reporting mechanisms (scaled scores and performance descriptors). One of the first of the reforms, the removal of national curriculum levels, was part of the policy direction of increasing school autonomy and seeking to develop a

'school-led system.' Although welcoming the move to greater school autonomy, Russell Hobby of the NAHT has criticised the 'breakneck speed' of the changes to the assessment system (Hobby, 2016).

- 16. We recommend a robust evaluation of the impact of the assessment reforms on teaching and learning.** It is probably too early to draw conclusions about the longer-term impact of the assessment reforms on teaching and learning but a timely evaluation should be considered. This could also address the question of how best to introduce changes to the system in future to minimise the burdens and uncertainty faced by teachers about the operation and implications of the assessment system.

## **Logistics and delivery of the 'SATs'**

17. The current National Curriculum Tests undergo a robust process of development. There are four main stages: item (question) writing, trialling, standardisation and standard setting. Typically, test development starts with a test specification and an item (question) specification. These set out, among other things, the parts of the curriculum to be tested and define the technical characteristics of the test. Once test questions have been drafted, the next step is to informally trial, or pilot them. Feedback from this phase of the process is used to refine the test items, which are then submitted to professional review groups before formal trials are carried out on a large-scale sample. Rigorous analysis of trialling results is conducted and the best items are selected for the final version of the test. A standard setting process will determine the boundaries for the assessment. All these stages ensure the tests are valid, reliable and fit for purpose.

## **Training and support needed for teachers and senior leaders to design and implement effective assessment systems**

- 18. For any system of assessment to be effective teachers must understand the purposes of the key components of the system, as well as the practical skills of administering assessments, interpreting the results and providing feedback to pupils, parents / carers, colleagues and school leaders.** Initial teacher education clearly needs to address these issues but cannot devote the time required to develop this expertise. Much of the professional development needs to take place once teachers are in post. Moderation procedures of teacher assessment judgements or teacher marking can provide one of the best means of developing assessment expertise.
- 19. NFER believes that a shared understanding of assessment is inextricably linked with teachers' professional development.** Whatever the future shape of assessment, one thing is certain: we can best support pupil learning and progression through a commitment to supporting the professional understanding of educational assessment. In order to introduce and maintain rigorous and consistent approaches to assessment in schools, novice and established teachers need to develop a culture and discourse of high quality assessment throughout their careers. The role of professional associations, advisers and schools should be to support this development, so that a shared understanding of assessment can become embedded in classroom practice.

20. **There is a need for greater assessment data literacy amongst primary school teachers.** In the context of the key stage national curriculum assessments, greater support and training is also required to help teachers understand the new scaled score metric used to report the outcomes of the tests. This issue of the need for greater assessment literacy was identified in the Carter review into Initial Teacher Training in 2015 and is relevant not just to new teachers but to experienced teachers as they come to terms with different reporting outcomes since the removal of national curriculum levels. **Consideration should be given to providing funding for a programme of support and training for assessment data literacy.**

## References

Bew, P. (2011). *Independent Review of Key Stage 2 Testing, Assessment and Accountability. Final Report*. London: DfE [online]. Available: <http://www.educationengland.org.uk/documents/pdfs/2011-bew-report-ks2tests.pdf> [26 October, 2016].

Burge, B., Lenekit, J. and Sizmur, J. (2015). *PISA in Practice – Cognitive Activation in Maths: How to use it in the classroom*. Slough: NFER

Burgess, S., Wilson, D. and Worth, J. (2013). 'A natural experiment in school accountability: the impact of school performance information on pupil progress'. *Journal of Public Economics*, 106, 57-67.

Campbell, T. (2013). *Stereotyped at Seven? Biases in Teacher Judgements of Pupils' Ability and Attainment* (CLS Working Paper 2013/8). London: IoE [online]. Available: <http://www.cls.ioe.ac.uk/shared/get-file.ashx?itemtype=document&id=1715%20> [26 October, 2016].

Carter, A. (2015). *Carter Review of Initial Teacher Training (ITT)*. London: DfE [online]. Available: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/399957/Carter\\_Review.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/399957/Carter_Review.pdf) [26 October, 2016].

Department for Education, (2016) Reducing teacher workload: Data Management Review Group report  
<https://www.gov.uk/government/publications/reducing-teacher-workload-data-management-review-group-report>

Department for Education, Commission on Assessment without Levels (2015). *Final Report of the Commission on Assessment with Levels*. London: DfE [online]. Available: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/483058/Commission\\_on\\_Assessment\\_Without\\_Levels\\_-\\_report.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/483058/Commission_on_Assessment_Without_Levels_-_report.pdf) [26 October, 2016].

Department for Education and Timpson, E. (2016) *Funding Boost for Schools Helping Pupils Develop Character* (Press Release). London: DfE [online]. Available: <https://www.gov.uk/government/news/funding-boost-for-schools-helping-pupils-develop-character>) [26 October, 2016].

Department for Education (2015). *2016 key stage 2 science sampling test framework: national curriculum tests from 2016*. London: DfE [online]. Available: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/439614/2016\\_KS2\\_Sciencesampling\\_framework\\_PDFA.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/439614/2016_KS2_Sciencesampling_framework_PDFA.pdf) [26 October, 2016].

Hobby, R. (2016) Speech to NAHT conference [online]. Available: <http://www.naht.org.uk/welcome/news-and-media/key-topics/assessment/russell-hobbys-speech-to-naht-conference/> [26 October, 2016].

Lynch, S., Bamforth, H. and Sims, D. (2015). *Reception Baseline Research: Views of Teachers, School Leaders, Parents and Carers (DfE Research Report 409)*. London: DfE



Maughan, S., Cooper, L. and Benefield, P. (2009). *School Report Cards: Review of the Use of School Report Cards Internationally*. Slough: NFER.

NFER. (2013). *Consultation Response: Primary Assessment and Accountability under the New National Curriculum*. Slough: NFER [online]. Available: <https://www.nfer.ac.uk/publications/99941/> [26 October, 2016].

Newton, P.E. (2007). 'Clarifying the purposes of educational assessment', *Assessment in Education: Principles, Policy & Practice*, **14**, 2, 149-170. [online]. Available: <http://dx.doi.org/10.1080/09695940701478321> [26 October, 2016].

Ofsted (2016). *Foreign Languages and Science Provision in Primary Schools*. London: Ofsted [online]. Available: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/524559/Foreign\\_languages\\_and\\_science\\_in\\_primary\\_schools.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/524559/Foreign_languages_and_science_in_primary_schools.pdf) [26 October, 2016].

Standards and Testing Agency (2015). *Science Sampling Test Framework. National Curriculum Test from 2016*. London: STA [online]. Available: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/439614/2016\\_KS2\\_Sciencesampling\\_framework\\_PDFA.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/439614/2016_KS2_Sciencesampling_framework_PDFA.pdf) [26 October, 2016].

Tisi, J., Whitehouse, G., Maughan, S. and Burdett, N. (2013). *A Review on Marking Reliability Research (Report for Ofqual)* Slough: NFER [online]. Available: <https://www.nfer.ac.uk/publications/MARK01> [26 October, 2016].

Welsh Government. (2013). *A Rapid Evidence Assessment on the Impact of Curriculum and Assessment Arrangements within High Performing Countries* [online]. Available: <http://gov.wales/statistics-and-research/rapid-evidence-assessment-impact-curriculum-arrangements-high-performing-countries/?lang=en> [26 October, 2016].

## Annex 1 : Introduction to NFER

21. **The National Foundation for Educational Research (NFER) is a leading independent provider of robust and innovative research, assessments and related services.** As a charity, impact is at the heart of everything we do. We draw on evidence to provide insights into education from early years to higher education, to help improve education for all children and young people.
22. **During the last year we worked with around half of schools in England,** including schools that participated in our assessment trialling and sampling work. Our collaboration with schools includes research and evaluation, test development and trialling for national assessment programmes (working with organisations including DfE and the Welsh Government) and providing products and research evidence to schools.
23. **We also work extensively on international surveys,** administering and analysing the Programme for International Student Assessment (PISA), the Progress in International Reading Literacy Study (PIRLS) and the Trends in International Mathematics and Science Study (TIMSS) assessments in the four countries of the UK. In addition, we have reported on how evidence from international assessments can be used to support teachers' work in the classroom (e.g. PISA in Practice – Cognitive Activation in Maths: How to use it in the classroom (2015)).
24. **Our range of assessments and related services include NFER's own suite of robust primary phase tests for schools.** These are widely known and recognised for providing teachers and schools with targeted information on their learners' performance and for providing high-quality feedback to learners and schools. We were one of the three approved suppliers of Reception Baseline Assessments in 2015/16 and, in response to positive user feedback, we continue to offer the NFER Reception Baseline Assessment to meet continuing demand from schools.