



NFER Thinks

What the evidence tells us

**Evidence for
Excellence in
Education**

Performance descriptors risk undermining the benefits of teacher assessment



Teacher assessment, when it is done well, can be a powerful and effective way of enabling pupils to make progress.

In October 2014, in the latest stage of government reforms to the assessment and accountability arrangements for primary schools, a consultation on performance descriptors for use in statutory teacher assessment was announced (Department for Education, 2014a).

This *NFER Thinks* examines the implications of the proposed performance descriptors, and argues that without revisions they risk undermining the benefits of teacher assessment.

Background

In September 2014, alongside the introduction of the new National Curriculum and the removal of National Curriculum levels (Department for Education, 2013a; Brill and Twist, 2013), the Government published new accountability and assessment arrangements (Department for Education, 2014b). Central to these changes was the principle that both progress and attainment are crucial as measures of accountability. For primary schools they include the introduction of more challenging National Curriculum tests in 2016 with scaled scores, performance descriptors to inform teacher assessment, and an improved moderation system.

The latest stage in these reforms was marked in the autumn with a consultation on performance descriptors to be used at the end of Key Stages 1 and 2 (Department for Education, 2014a). Performance descriptors describe what pupils know and are able to do in relation to a 'national standard'. They reflect the higher standards pupils are expected to achieve at the end of Key Stage 2 (KS2), linked to the more challenging National Curriculum.

The stated aim of the performance descriptors is to 'underpin' teacher assessment (Department for Education, 2014a). However, their introduction, in the current proposed form, may have unintended adverse consequences for teacher assessment.

Context: teacher assessment

Teacher assessment is used both formatively, to inform day-to-day teaching and learning, and summatively, to summarise pupil performance at educational transition points. When it is done well, it can be a powerful and effective way of enabling pupils to make progress. It can pinpoint a pupil's strengths and weaknesses; indicate the next steps necessary for progress in a particular area of learning; facilitate the setting of specific and meaningful individual targets; and can inform the planning of teaching and learning opportunities, including the differentiated learning needs of groups and individuals (Black and William, 1998; Mansell *et al.*, 2009).

In England, teacher assessment has a key summative role. As illustrated in **Figure 1**, it complements the National Curriculum tests at the end of Key Stages 1 and 2 (KS1 and KS2) by providing a more rounded picture of individual attainment. The National Curriculum tests serve an important purpose, capturing a snapshot of attainment in the core subjects at a particular point in time and facilitating comparison of individual and group performance with national standards. However, they have a relatively narrow focus. Teacher assessment has the potential to cover all aspects of the National Curriculum, including those that can be assessed only by means of practical and oral class work, and draws on a wider evidence base, using a range of sources including observation, questioning, written work, informal and formal testing and self and peer evaluations.

There is persistent debate about the reliability of teacher assessment used summatively. According to the evidence submitted to the National Association of Head Teachers (NAHT) Commission on assessment, the use of external testing over many years has 'eroded both public and professional confidence' in teacher assessment (NAHT, 2014).

We argue that there are three main reasons for this conclusion:

- The inclusion of test rather than teacher assessment outcomes in performance tables suggests greater importance
- The use of sublevels implies a spurious accuracy in teacher assessment when there are, in fact, no nationally agreed definitions of performance at any sublevel
- Evidence of bias.

In relation to this third point, it is recognised that teacher assessment can be unreliable or subject to bias (Campbell, 2013; Harlen, 2004), although the Harlen review also found that the reliability of teacher assessment was highly dependent on context. For example, reliable teacher assessment was supported when progress was assessed against a framework of finely specified criteria describing different levels of achievement (Rowe and Hill, 1996). Moderation was found to be beneficial not only in supporting assessment judgements, but also in developing teachers' understanding of their expectations of pupil performance.

The high-stakes accountability context within the English educational system also has an impact on teacher assessment. For example, pupils' attainment in KS1 teacher assessments is more likely to be judged higher in infant or first schools compared to all-through primary schools (Ofsted, 2013), which is likely to be influenced by the greater importance of these outcomes for measuring infant schools' effectiveness (Withey and Turner, 2015).

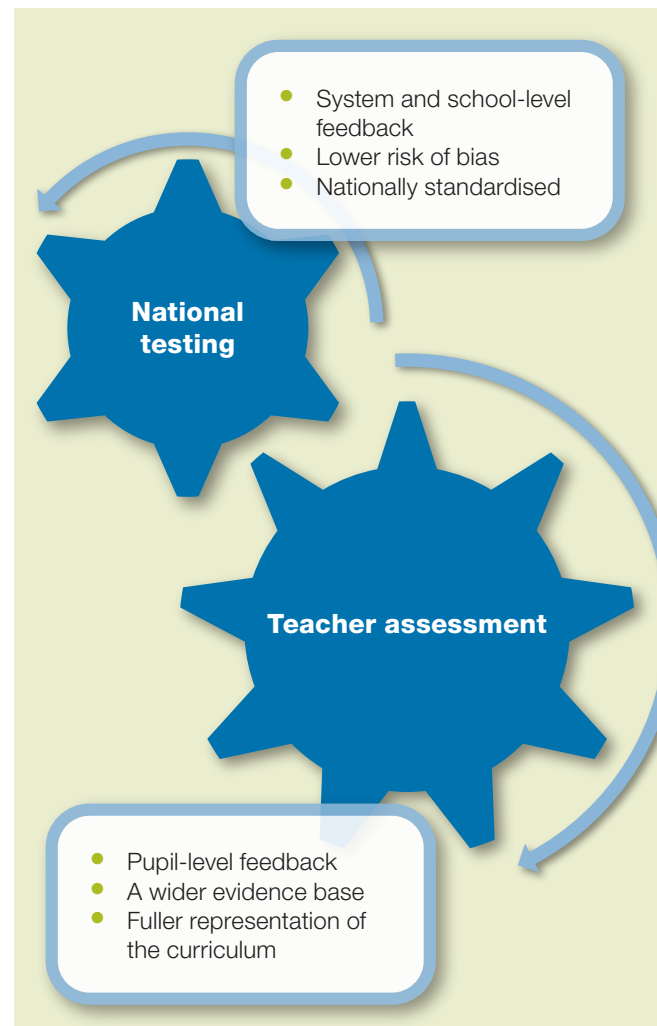


Figure 1: The complementary roles of teacher assessment and national testing

Performance descriptors

Within this context, the performance descriptors are being introduced to support teacher assessment at the end of Key Stages 1 and 2. For each subject and key stage, the proposed number of performance descriptors is dependent on whether teacher assessment is used to hold schools to account, i.e. if summative teacher assessment contributes to the 'floor standard' (the minimum requirements schools are expected to achieve).

Key stage	Subject	Performance descriptors
KS1	Reading*	mastery standard
	Writing*	national standard
	Mathematics*	working towards national standard below national standard
KS1	Science	working at national standard
KS2	Reading	working at national standard
	Mathematics	
	Science	
KS2	Writing*	mastery standard
		above national standard
		national standard
		working towards national standard below national standard

* summative teacher assessment contributes to floor standard

Unintended consequences

The performance descriptors describe what pupils are expected to have achieved at the end of each key stage. Where only one performance descriptor is provided, it is assumed that teacher assessment will also be informed by the outcomes of the National Curriculum tests. However, the proposed structure may make teacher assessment of pupil attainment more problematic, and place additional burdens on teachers and schools.

- **Single performance descriptors risk undermining renewed emphasis on pupil progress.** The introduction of single performance descriptors for some subjects/stages seems in conflict with other reforms where progress is given prominence. For example, in secondary schools the GCSE A*–C floor standard is to be replaced with a Progress 8 measure, which depends on the progress of all students and removes the incentive for schools to focus their attention primarily on students near the C threshold. However, single performance descriptors (i.e. achieved/not achieved) risk conveying a message that attainment takes precedence over progress, focusing the attention of schools (and the media) on whether pupils are reaching the national standard rather than ensuring that all pupils are making at least satisfactory progress and that disadvantaged pupils are closing the gap.

- **The new arrangements could be less well understood by parents.** One of the reasons for the removal of National Curriculum levels was that '... this system is complicated and difficult to understand, especially for parents' (Department for Education, 2013a), and yet inconsistencies in the number of performance descriptors could create more confusion, not less. The Department for Education consultation document states that Ofsted will continue to expect schools to monitor the progress of each pupil and to communicate such progress to pupils, parents and governors. Single performance descriptors will not support teachers in doing this. Parents and other stakeholders will need KS2 outcomes to be explained and put into context, particularly for pupils below or well above the national standard. Over time, parents familiar with a range of performance descriptors at the end of KS1 will come to expect something similar at the end of KS2.
- **The lack of a full range of performance descriptors risks placing additional burdens on schools.** According to the Teachers' Standards guidance (Department for Education, 2011), teachers must:
 - *set goals that stretch and challenge pupils of all backgrounds, abilities and dispositions*
 - *be accountable for pupils' attainment, progress and outcomes.*

To do this, teachers will have to work out their own expectations of the attainment of pupils significantly above or below the national standard. This introduces an unnecessary burden at a time when they are still familiarising themselves with the requirements of the new National Curriculum; and learning to live without levels. It will take time for teachers to develop a comprehensive understanding of the new national standard expected at the end of each key stage and for them to work backwards and forwards within a key stage to ensure pupils are on track each year to make satisfactory progress or better. However, particularly for pupils above the national standard, teachers will have to do this without any support as to what end of KS2 performance might be expected nationally. In the short term this may lead to the more able pupils being insufficiently challenged. If teachers are required to make 'effective and consistent assessments' (Department for Education, 2014a, p.5), they either need to be given an 'agreed set of criteria and standards' or they need 'a more sophisticated infrastructure of guidance, training, support and cross-checking' (Mansell *et al.*, 2009).

- **The performance descriptors risk promoting a mechanistic tick-list approach to teacher assessment.** In contrast to the large number of bullets or statements within the descriptors, there is a lack of guidance about how teachers should make their decisions. There is, for example, ambiguity inherent in the stem used throughout: 'A pupil is able to demonstrate sufficient evidence of the following.' This in itself requires judgement about what constitutes 'sufficient'. The only apparent guidance is that 'pupils must demonstrate the majority of the elements described' (Department for Education, 2014a, p.6); this could be interpreted anywhere between 51 per cent and 99 per cent. If unchanged, in no time at all, algorithms will be available to translate mechanistic 'tick-list' recording into a percentage achieved or some other indication of whether the achievement constitutes 'sufficient' evidence of the descriptor. These will have the same difficulties as the sublevels developed as part of the previous system, which had no agreed national definition, but where multiple local descriptors were developed and used. The performance descriptors need to focus on the key elements of the subject at that stage that underpin future progress.

Although schools may welcome the more detailed nature of the performance descriptors (compared to the lean National Curriculum level descriptors), there is still much work to be done in refining them. At present, the descriptors appear to reflect what pupils will be taught, which is not the same as what pupils need to know and understand in order to progress in the subject. There is little recognition that in each subject there are key underpinning elements upon which subsequent learning is built. In order to support teacher assessment more fully, these need to be identified and to be the focus of the descriptors of achievement.

What needs to be done?

We recommend that the Department for Education (via the Standards and Testing Agency) should:

- **provide a range of KS2 performance descriptors both above and below the national standard for reading and mathematics (and science).** This would reinforce the importance of measuring progress as well as attainment at the end of KS2 and support teachers in setting realistic expectations and recognising the achievements of all pupils.
- **reduce and refine the performance descriptors to encapsulate only the key elements that pupils need to know and understand in order to make further progress in the subjects.** This would ensure that teachers focus attention on assessing important milestones in each subject rather than monitoring everything that has been taught. This clarification would also enable better monitoring by school leaders and Ofsted.
- **develop exemplification (in the form of moderated samples of pupils' work) to cover the range of performance seen at the end of each key stage, not just exemplification of the national standard.** This should be developed in conjunction with schools and school groups and would support consistency in teacher assessment.

- **build capacity in teacher assessment and improve the reliability of teachers' professional judgements, by giving moderation a high priority.**

This would further support consistency of teacher judgements of pupil progress and attainment. The recently published Carter review (Carter, 2015) into initial teacher training (ITT) emphasised the need to enhance the training in assessment theory and practice in ITT. We would argue this is a requirement system-wide.

Conclusion

Teacher assessment is an important component of a teacher's toolkit. However, as with any tool, it can be deployed to good effect or bad. The new performance descriptors, as proposed, risk blunting its effectiveness, and so require careful consideration.

However, they do also represent an opportunity to enhance teacher assessment. Exemplification of the performance descriptors and moderation of their use in teacher assessment would help to embed understanding of the national standards and support teachers' professional development in assessment. And, if they are modified according to our recommendations, they have the potential to support effective good quality teacher assessment; provide useful communication tools for pupils, parents and other stakeholders; demonstrate that attainment and progress are equally important; and – ultimately – improve the quality of teaching and learning.

References

- Black, P. and William, D. (1998). *Inside the Black Box: Raising Standards Through Classroom Assessment*. London: King's College.
- Brill, F. and Twist, L. (2013). *Where Have All the Levels Gone? The Importance of a Shared Understanding of Assessment at a Time of Major Policy Change* (NFER Thinks: What the Evidence Tells Us). Slough: NFER [online]. Available: www.nfer.ac.uk/publications/99940/99940.pdf [28 January 2015].
- Campbell, T. (2013). *Stereotyped at Seven? Biases in Teacher Judgements of Pupils' Ability and Attainment* (CLS Working Paper 2013/8). London: University of London, Institute of Education, Centre for Longitudinal Studies [online]. Available: www.cls.ioe.ac.uk/shared/get-file.ashx?itemtype=document&id=1715 [28 January 2015].
- Carter, A. (2015). *Carter Review of Initial Teacher Training (ITT)*. London: Department for Education [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/399957/Carter_Review.pdf [2 February 2015].
- Department for Education (2011). *Teachers' Standards: Guidance for School Leaders, School Staff and Governing Bodies*. London: Department for Education [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/301107/Teachers__Standards.pdf [28 January 2015].
- Department for Education (2013a). *Assessing Without Levels* [online]. Available: <http://webarchive.nationalarchives.gov.uk/20130123124929/http://www.education.gov.uk/schools/teachingandlearning/curriculum/nationalcurriculum2014/a00225864/assessing-without-levels> [3 February 2015].
- Department for Education and Morgan, N. (2013b). *Reforming Qualifications and the Curriculum to Better Prepare Pupils for Life After School* [online]. Available: <https://www.gov.uk/government/policies/reforming-qualifications-and-the-curriculum-to-better-prepare-pupils-for-life-after-school> [28 January 2015].
- Department for Education (2014a). *Performance Descriptors for Use in Key Stage 1 and 2 Statutory Teacher Assessment for 2015/16* [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/368298/KS1-KS2_Performance_descriptors_consultation.pdf [28 January 2015].
- Department for Education (2014b). *Reforming Assessment and Accountability for Primary Schools: Government Response to Consultation on Primary School Assessment and Accountability*. London: Department for Education [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/297595/Primary_Accountability_and_Assessment_Consultation_Response.pdf [28 January 2015].
- Harlen, W. (2004). 'A systematic review of the evidence of reliability and validity of assessment by teachers used for summative purposes.' In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London [online]. Available: http://eppi.ioe.ac.uk/eppi/Evidence/EPPI_reviews/Assessment/Review3/ass_rv3.pdf [28 January 2015].
- Mansell, W., James, M. and the Assessment Reform Group (2009). *Assessment in Schools. Fit for Purpose? A Commentary by the Teaching and Learning Research Programme*. London: Economic and Social Research Council, Teaching and Learning Research Programme [online]. Available: www.tlrp.org/pub/documents/assessment.pdf [28 January 2015].
- NAHT (2014). *Report of the NAHT Commission on Assessment*. Haywards Heath: NAHT [online]. Available: www.naht.org.uk/assets/assessment-commission-report.pdf [28 January 2015].
- NFER, 2013. *Practical Guide to the 2014 National Curriculum, Volumes 1–3*. Slough: NFER.
- Ofsted (2013). *Ofsted Annual Report 2012/13*. London: Ofsted [online]. Available: <https://www.gov.uk/government/collections/ofsted-annual-report-201213> [28 January 2015].
- Rowe, K.J. and Hill, P.W. (1996). 'Assessing, recording and reporting students' educational progress: the case for 'subject profiles', *Assessment in Education*, **3**, 3, 309–352.
- Stanley, G., MacCann, R., Gardner, J., Reynolds, L. and Wild, I. (2009). *Review of Teacher Assessment: Evidence of What Works Best and Issues for Development*. London: QCA [online]. Available: http://oucea.education.ox.ac.uk/wordpress/wp-content/uploads/2011/01/2009_03-Review_of_teacher_assessment-QCA.pdf [28 January 2015].
- Withey, P. and Turner, S. (2015). 'The analysis of SATS results as a measure of pupil progress across educational transitions', *Educational Review*, **67**, 1, 23–34.

NFER

NFER is a charity with a reputation worldwide for providing independent and robust evidence for excellence in education and children's services. Our aim is to improve education and learning, and hence the lives of learners, by researching what is happening now. Our authoritative insights inform policy and practice, offering a unique perspective on today's educational challenges.

We work with an extensive network of organisations, all genuinely interested in making a difference to education and learners. Any surplus generated is reinvested in research projects to continue our work to improve the life chances of all learners, from early years through to higher education.

Catherine Kirkup



Catherine has extensive experience managing and directing test development projects and other assessment research. She is a Practitioner of the Association for Educational Assessment-Europe and a Fellow of the Chartered Institute of Educational Assessors, and co-author of a number of authoritative guides on assessment, including NFER's *Practical Guide to the 2014 National Curriculum, Volumes 1–3* (NFER, 2013).

Liz Twist



Liz oversees all assessment research and development projects at NFER. She has directed a large number of assessment development projects, primarily in the area of literacy, and many on a national scale. These included the Key Stage 2 tests in English between 1997 and 2012, taken by all 11-year-olds in England. She currently leads the development, trialling and delivery of national tests in reading and numeracy for the Welsh Government.

Published February 2015

© National Foundation for Educational Research
Registered charity no 313392

978-1-910008-50-8

How to cite this publication:

Kirkup, C. and Twist, L. (2015). *Performance Descriptors Risk Undermining the Benefits of Teacher Assessment* (NFER Thinks: What the Evidence Tells Us). Slough: NFER.

**National Foundation for
Educational Research**
The Mere, Upton Park
Slough, Berks SL1 2DQ

T: 01753 574123
F: 01753 691632
E: enquiries@nfer.ac.uk
www.nfer.ac.uk

