
The Assessment System is Unsustainable: how can we make it better?

GEMMA MOSS

ABSTRACT This article evaluates the current state of play in terms of policy development in response to the testing crisis.

The UK government's new consultation *Primary Assessment in England* (March-June 2017) is to be welcomed. The key question is: will it go far enough? The answer to that in large part depends upon how parents, teachers and academics react.

What's Driving the Consultation?

The terms of the consultation are worth quoting from the Department for Education (DfE) website:

Recognising the scale of the changes that we have asked primary schools to deal with, this consultation represents a significant step towards establishing a settled, stable primary assessment system that is trusted by teachers and parents. (DfE, 2017)

The phrase 'a settled, stable primary assessment system that is trusted by teachers and parents' is well chosen, for at the moment we do not have that. Full credit to the Secretary of State for Education for setting this as the consultation's main goal.

What Has the Government Decided Will Change?

Some changes are already under way. The unhappy debacle over primary testing last year left many children facing questions they found too difficult to

answer, based on an ever-narrower subset of skills that do not reflect the full range of competencies they should develop in a rich and purposeful curriculum. The announcement from the DfE that the Standards and Testing Agency has changed the type and difficulty of the questions at the start of each test 'to ensure children are not discouraged' (Greening, 2017) goes a little way towards addressing this. But a broader review of the impact that such poorly conceived primary assessment instruments are now having on the wider curriculum is also badly needed.

This is not quite what the consultation sets in motion. It is proposing that:

- Key Stage 1 (KS1) tests become optional and do not form part of the progress measures used to hold schools to account;
- any baseline tests put into the reception stage for accountability purposes will not be publicly reported at school level;
- the current approach to the statutory assessment of writing at KS2 will be revised, with the value of teacher assessment given more status and the trialling of peer-to-peer moderation to see if this can usefully bring matters more fully back into teachers' hands.

All of this signals a move away from a tick-box testing culture based on closed questions and a choice between right or wrong answers capable of being marked by a machine. Although the controversial SPaG test (spelling, punctuation and grammar) at KS2 and the Phonics Check at KS1 remain in place, the questions raised in the consultation are moving the assessment debate in the right direction. However, it is crucial that interested parties – teachers, parents, educational researchers – continue to argue for change.

What Else Needs to Change?

The current assessment system confuses collecting test data in order to hold schools to account with using assessment data to identify what children already know and understand in ways that further their learning.

These two purposes are quite distinct, even if talk of raising standards and measuring pupil progress might treat them as the same. England is an outlier internationally in the amount of tests and the severity of the penalties schools face if they do not perform well. The negative consequences of such an approach are well understood. As the Organisation for Economic Co-operation and Development (OECD) comments in a review of assessment practices of 28 countries:

Evaluation and assessment should serve and advance educational goals and student learning objectives... Because of their role in providing accountability, evaluation and assessment systems can distort how and what students are taught ... [I]f teachers are judged largely on results from standardised student tests, they may 'teach to the test', focusing solely on skills that are tested and giving less

attention to students' wider developmental and educational needs.
(OECD, 2013)

On Baseline, Benchmarking and Measuring Pupil Progress

The consultation invites responses on the use of assessment data to hold schools to account for the progress pupils make. This is framed as a choice between using tests in reception or at KS1 to measure pupil progress to KS2. It is suggested that testing in reception will give schools more credit for the work that they do prior to Year 2. However, it is worth pausing over the assumption that the right conclusions can be drawn about the value that schools add to pupils from benchmarking and measuring pupil progress in this way.

Some Questions to Think About

1. Is the government right in assuming that the progress pupils make between baseline and KS2 outcomes is an adequate means of judging one school's performance against another's? The simple answer is no. The school effectiveness literature shows that while prior attainment is the most important predictor of final attainment, student demographic and socioeconomic characteristics also matter hugely. Unless these factors are taken into account, progress measures based on national norms will be biased in favour of schools with middle-class intakes. Those schools serving the most disadvantaged areas will continue to be at much greater risk of the performance data being used to trigger hostile interventions, including forced academisation and the loss of their Head. (See Leckie & Goldstein, 2017, for a more detailed discussion of the limits to the statistical models deployed to date.)

2. Are the available methods for tracking pupil progress really adequate for the purposes to which they are being put? The most common approach is to average the progress pupils make from particular starting points and through subsequent tests to produce smooth and linear lines of progression against which any individual pupil's performance will be judged. But such models bear little relationship to the actual paths that any particular child may take. These show much more variation than the models allow. Most children do not make average progress each year, but on the contrary 'experience periods of both slower and more rapid progress' (Education Datalab, 2015). We do children a profound disservice if we insist they must learn at the same pace, or that those at further distance from the national standard must make up ground most rapidly.

3. How confident can we be that any baseline tests developed for the early years will accurately assess the abilities they seek to capture? The government abandoned previous attempts to let schools commission baseline tests from a range of different suppliers – Early Excellence, the Centre for Evaluation and Monitoring (CEM), and the National Foundation for Educational Research (NFER) – because the different tests gave results which were not strictly comparable. These difficulties have not gone away. Some of these same test developers have put on record that it is very difficult to make accurate assessments of EAL

(English as an additional language) pupils' abilities when they enter school (Tymms et al, 2014). This contradicts the claim made in the consultation that:

it is possible to create an assessment of reception age children which is suitable for that age group, sufficiently granular and well correlated with later outcomes such that it could be used as a baseline from which to assess progress. (DfE, 2017, p. 16)

The consultation should clarify how the validity of any test instruments will be assured. It should make clear the risks to fair comparison that stem from measurement error in assessing prior attainments at an early age. So far there has been little attempt in the English system to extensively trial primary assessment instruments ahead of their introduction or to transparently review whether they adequately capture the competencies they purport to measure (Standards and Testing Agency, 2012, p. 52). Further, whether they are developmentally appropriate for the age at which children sit the test has been ignored (Standards and Testing Agency, 2013, p. 9). This needs to change.

Changing Things for the Better

The consultation opens up some questions for consideration, but it doesn't go far enough. The high-stakes accountability and high school-autonomy system we currently labour under is not good for children. It fosters risk-averse management practices in schools, excessive reliance on data monitoring and too much teaching to the test.

The pressure on the government to change has happened through parental lobby, actions taken by the teacher unions, and the alliances that have formed over these issues – More than a Score; Better without Baseline; Let Our Kids be Kids. These alliances are encouraging parents, teachers and academics to:

- organise a local meeting with teachers and other parents to share ideas about the kind of schooling we really want;
- ask headteachers what they are doing to bring about a fairer assessment system that works first and foremost in the interests of the children in their school;
- write to MPs voicing their concerns.

Educational researchers have a particular responsibility to act for the public good.

We have a necessary role in informing public debate on what assessment data can justifiably be used for and what they can't. There is an increasing body of evidence on accountability and testing that demonstrates that the English assessment system is unsustainable, leads to misleading comparisons, and is a disproportionate waste of teachers' time and resources. Accountability purposes place too much weight on test data without publicly communicating the statistical uncertainties involved in measuring pupil progress in these ways

(Leckie & Goldstein, 2017). Making teachers as well as parents more aware of the uncertainties hiding in the statistical models is an important task.

It would be good to see this consultation as the first step back to the kind of education system that teachers, parents and children want. But the political path to a better settlement in education is not straightforward. All the more reason to get involved.

References

- Department for Education (DfE) (2017) Primary Assessment in England. https://consult.education.gov.uk/assessment-policy-and-development/primary-assessment/supporting_documents/Primary%20assessment%20in%20England.pdf (accessed 28 July 2017).
- Education Datalab (2015) *Seven Things you Might Not Know about Our Schools*. <http://www.educationdatalab.org.uk/wp-content/uploads/2016/02/EduDataLab-7things.pdf> (accessed 28 July 2017).
- Greening, J (2017) Government to Consult on Reforms to Primary Assessment. <https://www.gov.uk/government/news/government-to-consult-on-reforms-to-primary-assessment-system>
- Leckie, G. & Goldstein, G. (2017) The Evolution of School League Tables in England 1992-2016: 'contextual value-added', 'expected progress' and 'progress 8', *British Educational Research Journal* 43(2), 193-212. <https://doi.org/10.1002/berj.3264>
- Organisation for Economic Co-operation and Development (OECD) (2013) Synergies for Better Learning: an international perspective on evaluation and assessment. <http://www.oecd.org/edu/school/synergies-for-better-learning.htm#1> (accessed 19 May 2017).
- Standards and Testing Agency (2012) *Year 1 Phonics Screening Check. Pilot 2011 – Technical Report*. London: DfE.
- Standards and Testing Agency (2013) *English Grammar, Punctuation and Spelling. 2013 Technical Report*. London: DfE.
- Tymms, P., Merrell, C., Hawker, D. & Nicholson, F. (2014) *Performance Indicators in Primary Schools: a comparison of performance on entry to school and the progress made in the first year in England and four other jurisdictions. Research report*. London: DfE.

GEMMA MOSS is Professor of Literacy and director of the International Literacy Centre at the UCL Institute of Education, London. She has substantial experience as a teacher and researcher, and was president of BERA 2015-17. Correspondence: gemma.moss@ucl.ac.uk

