

Response ID ANON-6FQ1-TW38-M

Submitted to **Primary assessment in England**

Submitted on **2017-06-22 16:53:03**

Ministerial Foreword

Introduction

1 What is your name?

Name:

Gemma Moss

2 What is your email address?

Email:

gemma.moss@ucl.ac.uk

3 What is your organisation?

Organisation:

British Educational Research Association

4 Which of these categories best describes your role?

Categories:

Academic/Researcher

5 Would you like us to keep your responses confidential?

No

Reason for confidentiality:

Current system: statutory assessment in key stages 1 and 2

Preparing children to succeed at school

6 The EYFSP measures a child's development against the ELGs set out in the EYFS statutory framework. Should the profile be improved to better assess a child's knowledge, skill, understanding and level of development at the end of the early years? If so, please describe which elements could be added, removed or modified.

Please write your response in this box.:

The Department needs to be very clear on the uses to which it wishes to put the EYFSP data, and clarify what these are before introducing any changes. Some of the arguments put forward in the section on The early years foundation stage and profile do not stand up to scrutiny. In particular the observation that "the gap in reading between disadvantaged children and their peers, is wider in England relative to other developed countries." is not borne out in the reference cited. The Sutton Trust report on International Inequalities in fact records that "England is roughly in the middle of the pack (five countries have larger gaps, two have similar ones, and four have smaller ones)." Overstating the scale of the problems that England faces does not help make the Department's case.

One of the purposes of retaining the EYFSP, highlighted in the consultation document, is to encourage the development of best practice within the early years sector. This is an important aim. Researchers using the EYFSP to explore educational attainment have identified some key areas of development in vocabulary and in self-regulation which may hold the key to narrowing the attainment gap (See Bradbury et al, 2012). This is important information from exploring the data at scale which can be usefully fed back from researchers to practitioners. But this is very different from using school level data to hold individual institutions to account for the progress children make through school, on the assumption that we know exactly how each child with this profile will progress from here. On the contrary, we know there is considerable variation in the particular trajectories to children's learning in the early years. Not all children follow similar routes. From the practitioner perspective the value of the EYFSP may therefore lie much more precisely in the information it yields about any one child, enabling professional judgement to be exercised over how to address their particular needs in this context.

This consultation needs to give full consideration to the diversity of views it elicits from different parties. Ideally this should be a first step in building a real consensus between practitioners, researchers and policymakers on how the EYFSP can best be used to enrich early years education for all. Certainly BERA would like to see more dialogue between quantitative researchers working on child development and practitioners who can take those insights into the early years setting and use them to develop new pedagogic strategies in that light. Government has a role in building such productive research partnerships over the medium to long term. Promoting best practice in this way is very different from holding schools individually to account.

7 The EYFSP currently provides an assessment as to whether a child is 'emerging, expecting or exceeding' the level of development in each ELG. Is this categorisation the right approach? Is it the right approach for children with SEND?

Please write your response in this box.:

BERA agrees with TACTYC's submission that these categorisations are very broad brush, and that "The existing practice of reporting results for the whole cohort with no adjustment for age is damaging to younger children in the group. Too many summer-born children, particularly boys, are misdiagnosed as having special needs when their scores are lower for developmental reasons and through relative lack of experience. By the end of KS1, an August-born child is 90% more likely to be diagnosed with SEND than a September-born child. (Research Report DFE-RR017, Month of Birth and Education, Schools Analysis and Research Division, July 2010)"

The refusal to recognise that age in months impacts on development leads to a number of perverse consequences for summer-born children. Standardised tests for young children standardise responses by age for a reason. Norm-referencing (which in effect the three categories "emerging, expecting or exceeding", represent) without taking age in months into account is hard to justify, and as TACTYC's submission makes clear, does some children a profound disservice.

Early years foundation stage profile: workload

8 What steps could we take to reduce the workload and time burden on those involved in administering the EYFSP?

Please write your response in this box.:

Continue to emphasise that the primary purposes of the EYFSP are to enhance children's learning, and as an adjunct to the exercise of professional judgement in planning access to a full and rich curriculum.

Early years foundation stage profile: moderation

9 How could we improve the consistency and effectiveness of the EYFSP moderation process whilst reducing burdens?

Please write your response in this box.:

The best starting point for measuring progress in primary school

10 Any form of progress measure requires a starting point. Do you agree that it is best to move to a baseline assessment in reception to cover the time a child is in primary school (reception to key stage 2)? If you agree, then please tell us what you think the key characteristics of a baseline assessment in reception should be. If you do not agree, then please explain why.

Please write your response in this box.:

There is an increasing body of research evidence demonstrating that using progress measures to hold schools to account is fraught with difficulties and does little to address equity issues in the school system (See Leckie and Goldstein, 2017 for a detailed discussion on value added methodologies and an early review of the impact of Progress 8). Moving to a baseline assessment measure in reception is particularly problematic given the marked variation in the pace and sequence at which children acquire a range of literacy and numeracy skills under normal conditions. Indeed, normal development in countries which start formal schooling at age 7 includes learning to read and write at that age. There are no penalties associated with not already having done so.

By introducing such a baseline measure in reception the risk is that undue importance will be given to formal learning in the early years, detracting from the evidence that scaffolded play, cognitively enriching activities, and language-rich social interactions with staff and peers all help to alleviate the effects of social disadvantage and have a protective effect for those most at risk of poor educational outcomes (Moss and Washbrook, 2016; Siraj-Blatchford, 2013; Sylva et al., 2007). The unintended consequences might thus be precisely the reverse of the desired effect.

The intention to use baseline measures in reception "to give schools credit for the value that they add in the first 3 years" is unlikely to yield the desired result. On current data, correlations of baseline data with KS2 test scores are at most around 0.5, and rather smaller than this where subscales or components of tests are concerned. Pupil mobility is also an issue that remains unaddressed.

Our main concerns can be summarised as follows:

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, then 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 and 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" (Allen, 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 and sustain that additional effort for longer.

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 test 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 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, p16)

There are considerable 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 (STA, 2012, p52). This includes ignoring whether they are developmentally appropriate for the age at which children sit the test (STA, 2013, p9). Far more recognition of the uncertainties hiding in the statistical models is needed. At the moment too much reliance is being given to test data that cannot bear the weight of interpretation placed upon them.

11 If we were to introduce a reception baseline, at what point in the reception year do you think it should be administered? In particular, we are interested in the impact on schools, pupils and teaching of administering the assessment at different times.

Please write your response in this box.:

We do not think it appropriate to introduce a reception baseline. Timing in the year will do little to address the core problems.

12 Our view is that it would be difficult to change key stage 1 assessment in order that it could be used as the baseline for progress in the long term. If you disagree, what could be done to improve the key stage 1 assessments so that they would be sufficiently detailed, and trusted as a fair and robust baseline?

Please write your response in this box.:

We agree that measuring progress in ways that do not systematically advantage schools with middle class catchments and systematically disadvantage schools with working class catchments are not easy to find, even with the use of sophisticated statistical techniques. These issues are discussed in depth in Leckie and Goldstein, 2017.

The best starting point for measuring progress in primary school: interim years

13 If we were to introduce a new reception baseline measure, do you agree that we should continue to use key stage 1 teacher assessment data as the baseline for measuring progress in the interim years before a new measure was in place? If you disagree, what do you think we should use as the baseline instead?

Please write your response in this box.:

The role of key stage 1 statutory assessments

14 If a baseline assessment is introduced in reception, in the longer term, would you favour removing the statutory requirement for all-through primary schools to administer assessments at the end of key stage 1?

Please write your response in this box.:

In their review of assessment practices of 28 countries, the OECD comment that 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. .. if 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. (OECD, 2013)

There is plenty of evidence that the use of high stakes assessment in the English system is indeed leading to teaching to the test. Accordingly, we favour the monitoring of national standards in place of statutory assessments at the end of Key stage 1.

Monitoring national standards at key stage 1

Measuring progress in different types of school

15 If we were to introduce a reception baseline to enable the creation of reception to key stage 2 progress measures for all-through primaries, what would be the most effective accountability arrangements for infant, middle and junior schools' progress measures?

Please write your response in this box.:

There are considerable difficulties in holding individual schools to account for pupil progress, given pupil mobility. Many pupils change schools. So there are a number of issues that challenge the assumption that progress data can be satisfactorily used for accountability purposes.

A proportionate assessment system

16 Do you think that the department should remove the statutory obligation to carry out teacher assessment in English reading and mathematics at key stage 2, when only test data is used in performance measures?

No

Please write your response in this box.:

The real issue is with the weighting given to teacher assessments, and the reliability and validity of the test instruments at KS 2. We urgently need a review of a tick box testing culture, based on closed questions and a choice between right or wrong answers (Bew, 2011, p61). The main rationale put forward for this has been to make tests capable of being marked by a machine. This may alleviate teacher workload, but has had the unintended consequences of considerably narrowing assessment at KS 2.

Key stage 1 English grammar, punctuation and spelling test

17 Do you agree that the key stage 1 English grammar, punctuation and spelling test should remain non-statutory beyond the 2016 to 2017 academic year, with test papers available for teachers to use as they see fit?

Yes

Please write your response in this box.:

The Spag test urgently needs review.

Multiplication tables check

18 At what point in key stage 2 do you think the multiplication tables check should be administered? Please explain the basis for your views.

Not Answered

Please write your response in this box.:

19 How can we ensure that the multiplication tables check is implemented in a way that balances burdens on schools with benefit to pupils?

Please write your response in this box.:

Reducing burdens within the primary assessment system

20 Are there additional ways, in the context of the proposed statutory assessments, that the administration of statutory assessments in primary schools could be improved to reduce burdens?

Please write your response in this box.:

Improving end-of-key stage statutory teacher assessment

21 Do you agree that the statutory assessment of writing should afford teachers greater flexibility in determining a pupil's overall standard of attainment than is currently the case? Please give reasons for your answer.

Yes

Please write your response in this box.:

Best fit better reflects teacher judgement and teachers should be afforded greater flexibility in exercising their professional judgement.

Supporting and strengthening the assessment of English writing

22 Please give details of any robust alternative approaches to the assessment of English writing, which the Department for Education should explore.

Please write your response in this box.:

Alternative approaches to moderation

23 Please give details of any effective models of moderation or standardisation of teacher assessment that the Department for Education should explore.

Please write your response in this box.:

Equalities

24 Do you think that any of our proposals could have a disproportionate impact, positive or negative, on specific students, in particular those with 'relevant protected characteristics' (including disability, gender, race and religion or belief)? Please provide evidence to support your response.

Please write your response in this box.:

See our answer to question 10. "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, then 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 and Goldstein, 2017, for a more detailed discussion of the limits to the statistical models deployed to date.) "

25 How could any adverse impact be reduced and are there any ways we could better advance equality of opportunity? Please provide evidence to support your response.

Please write your response in this box.:

By weakening the link between high stakes assessment and accountability, in line with OECD advice.

References for this submission

Bew, P. (2011) Independent Review of Key Stage 2 Testing, Assessment and Accountability. Final Report. London: DfE

Leckie, G. and Goldstein, H. (2017), The evolution of school league tables in England 1992–2016: 'Contextual value-added', 'expected progress' and 'progress 8'. *British Educational Research Journal*. doi:10.1002/berj.3264

Moss, G., & Washbrook, L. (2016). Understanding the Gender Gap in Literacy and Language Development (01/2016). Bristol Working Papers in Education Series.

OECD (2013) *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*. Paris: OECD

Siraj-Blatchford, I., Taggart, B., Mayo, A., Sammons, P., Melhuish, E., & Sylva, K. (2013). The learning life course of at "risk" children aged 3-16: Perceptions of students and parents about "succeeding against the odds". *Scottish Educational Review*, 45(2), 5–17.

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