

Response ID ANON-6FQ1-TF5W-4

Submitted to **Primary assessment in England**

Submitted on **2017-06-21 15:46:31**

Ministerial Foreword

Introduction

1 What is your name?

Name:

Andrew Lambirth

2 What is your email address?

Email:

A.Lambirth@greenwich.ac.uk

3 What is your organisation?

Organisation:

United Kingdom Literacy Association

4 Which of these categories best describes your role?

Categories:

Subject association

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.:

UKLA is pleased that the consultation recognises the strengths of the Early Years Foundation Stage, and its profile-based, rounded system of assessment. However, UKLA strongly believes that assessment should support the curriculum and not dominate it. Therefore UKLA would not favour any 'improved' assessment of 'a child's knowledge, skill, understanding and level of development' that threatened to narrow children's learning experiences and thus adversely affect their development. Harlen (2014) shows clearly that assessment has a negative impact when what is assessed concerns only easily measured or tested aspects of learning. The pressure for higher scores forces teachers to focus their teaching on assessment content: assessment drives teaching. Teaching is reduced to training pupils to perform the specific tasks to be assessed. Thus the classroom interpretation of the curriculum is narrowed to the detriment of pupils' wider and deeper understanding and skill.

UKLA also maintains that assessment for formative purposes and for informing parents should be separated from assessment for monitoring purposes. As we argue in our answer to Question 15 below, the latter should be carried out on a sampling basis, to ensure wide coverage of different aspects of literacy learning, while reducing the load borne by each class and teacher.

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.:

No. UKLA would like to see a return of the original version of EYFSP (DfES 2007). This early version of the EYFSP honoured young learners as being competent co-constructors of knowledge within social contexts where there was an explicit understanding of child-centred, play-based experiential learning. In the 2007 version, Reception class teachers were required to undertake 'observational assessments' throughout the year and in the summer term, when the children had turned five years old, grade the children according to the Early Years Foundation Stage Profile (EYFSP) criteria. The Profile is based upon Reception teachers' 'best fit' judgements within prescribed criteria.

In the revised Profile (2013) there is a much sharper focus on literacy and maths, with substantially raised thresholds, making them harder to achieve. Following these revisions to the EYFSP in 2013 the figures for children's achievements dropped from 64% (DfE 2012) to 52% with significantly worse figures for summer-born children, only 30% of whom reached the GLD (DfE 2012). Thus a high proportion of young children were labelled as 'falling behind' (Bradbury 2013, 656). UKLA believes that this is an example of negative and inappropriate assessment. As has been shown (Roberts-Holmes, 2015), this approach to early year's

assessment decontextualises and pathologises individual children, teachers and schools and ignores the effects of structural inequality within society (Lingard et al., 2013, 552). It is particularly problematic for younger children and those with SEND.

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.:

In the process of assessment, there is an inevitable trade-off between thoroughness and economy of time, which means that classroom assessment is threatened by superficiality on the one hand and excessive workload on the other. The purpose of assessment at this early stage should be formative – and reporting to parents. Thorough assessment for formative purposes is likely to remain time-consuming, but the time spent would be clearly fruitful in terms of helping teachers to tailor their teaching and their children's learning to the children's specific needs, thus making both more productive. It would also help teachers to communicate more effectively with parents. Such assessment needs to take account of the children's interests and strengths, and what might help them forward.

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.:

Since assessment in the Early Years should be carried out for formative purposes, rather than for the purposes of monitoring, there should not be a need for extensive moderation beyond the walls of the individual primary school.

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.:

The UKLA feels very strongly that the reception class is not the place to make a baseline assessment against which the child's subsequent progress through primary school is recorded. The Association draws on the report by Bradbury and Roberts-Holmes (2016) to demonstrate why a baseline assessment for very young children would be totally inappropriate

The report shows that the majority of schools already used informal on-entry assessments to plan teaching and to identify children with particular needs; the Baseline Assessment was not seen as an improvement on these methods.

The report also points out that it is well known from socio-cultural research that children learn through sets of social relationships (Broadhead, 2006; Flear, 2010). It suggests that authentic, holistic and developmentally appropriate assessment, based upon teachers' observations over time in a range of contexts, makes visible what young children are capable of learning in supportive and collaborative relationships. The best forms of assessment aim to make children's learning stimulating, rich and successful.

A child's well-being and the characteristics of effective learning, such as resilience, perseverance and self-regulation learnt in the context of meaningful play, are seen to be more reliable predictors of later academic achievement (Bodrova & Leong, 2007; Whitebread & Bingham, 2012; Siraj-Blatchford & Kingston 2015) rather than 'short-term academic results' which may not last (TACTYC, 2015).

Bradbury and Roberts-Holmes (2016) show that teachers and school leaders have serious doubts as to the accuracy of the Baseline Assessment and its use in measuring progress. Only 7.7 per cent of respondents to their survey agreed the data was an 'accurate and fair way to assess children' because of the wide range of variables in the assessment process.

In addition, Bradbury et al (2016) report many teachers and school leaders doubt the use of measuring progress from Reception to Year 6 given the problems of assessing accurately at age four and the variability of children's patterns of progress and development.

In short, UKLA argues that assessment carried out in reception class should be for formative purposes, not to provide a baseline for subsequent assessment.

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.:

See above. UKLA does not recommend the introduction of baseline assessment into reception classes.

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.:

UKLA agrees that Key Stage 1 tests should not be used as a baseline for progress. Nevertheless, UKLA understands that if schools are to be judged fairly on the progress their pupils make, some baseline assessment is necessary. However this summative assessment, carried out at some point in KS1, for such purposes of accountability, should be clearly separate from formative assessment. The DfE should institute pilot studies to ascertain scope, foci and best timing for this summative assessment to yield a sound starting point for the identification of progress over time. It is essential that such an assessment yields more wide-ranging and nuanced information than a single number for reading or writing. Again, a sampling procedure is recommended for this exercise.

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.:

UKLA does not believe that a new reception baseline measure should be put in place. Please see above.

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.:

No. UKLA are concerned that if this requirement were removed, the only official means to assess children before the end of Key Stage 2 would be the Phonics check and the non-statutory English grammar, punctuation and spelling check, both of which UKLA believe to be totally inadequate and inappropriate forms of assessment.

The SPaG tests at years 2 and 6 reflect only a narrow section of the English curriculum, reducing it to a the naming linguistic elements with little regard to the function of language in use. They therefore give no indication of children's command of such forms to construct meaning through reading or writing. In addition, they are designed to test what children do not know rather than what they know and can do.

As research has indicated (Myhill et al. 2012), it would be more productive to teach the use of grammatical features of language as part of a wider approach to making meaning through writing. Grammar tests should be reframed to allow children to show what they know and can do in the context of more extended writing rather than 'gap-filling'.

UKLA considers that the Phonics Check has contributed very little of value to the reading assessment processes. Increasing scores on the Phonics Check have not been followed by improvements in scores on the Reading SAT for the same cohorts. Margaret Clark's (2015) work has highlighted serious flaws in the Check's validity as a robust assessment tool. The DfE's own commissioned report into the Phonics Check states:

'Analyses of pupils' literacy (reading and writing) scores in the national datasets over four years were not conclusive: there were no improvements in attainment or in progress that could be clearly attributed to the introduction of the check' (Walker et al., 2015, p. 8)

In addition, as Clark points out, the Phonics Check has no diagnostic value and includes no suggestion that methods of approaching learning to read other than phonics-based approaches may be appropriate for children who fail the test. UKLA find it regrettable that the DfE is not consulting on all aspects of the current assessment system in primary schools.

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.:

We do not believe the government should introduce a reception baseline for the reasons expressed above.

UKLA is concerned that the range of the curriculum in England suffers because of the focus on tests in English and Mathematics. As the government-commissioned survey of the impact of the Phonics Check has demonstrated, teaching to this test has narrowed the English Curriculum (Walker et al., 2015). The core subjects have been similarly impoverished.

Furthermore, there is no evidence that the tests and assessment processes currently being used improve practice in the target curriculum areas.

In addition, reports from teachers to UKLA members indicate the 'transaction costs' for doing the tests are disproportionate. The tests and the teacher assessment framework have resulted in greatly increased marking loads for teachers, taking teachers away from their main job in the classroom.

Schools experience pressure from regional Ofsted directors to account for their test results. As has already been stressed above, this can spread anxiety, for teachers as well as senior managers, and result in narrow 'teaching for the test' rather than the more fruitful approach of teaching the curriculum with knowledge of what the test involves.

This testing does not improve learning. Even within the narrow scope of their subject matter, the tests are not aimed at identifying future learning paths for pupils or teaching routes for teachers. Furthermore, the heavy focus on the written word is counter-productive: recent research into marking (Elliott et al., 2016) shows that 'The quality of existing evidence focused specifically on written marking is low' (p.5), calling into question the usefulness of current practices developed in response to tests and assessments. Written marking may give the impression of aiding pupil learning, but oral responses to children's work, delivered as the work is proceeding in the classroom may be more effective (Elliott et al., 2016).

Judgements made on numerical data from narrowly conceived tests do not give a full picture of pupil achievement. They are particularly problematic given the different demographics of schools and their catchment areas. Certainly, as indicated above, they do not help children move forward.

Because of their narrowness of focus, and the insufficient scrutiny of the validity and reliability of the data they yield, test results and teacher assessments based on the current performance descriptors should not be used as the only data in identifying schools in need of improvement.

As Harlen (2014) has pointed out:

'The arguments against using levels put forward by the government's expert group included concern that a level does not convey detailed information, although it can indicate progress from year to year. But now, instead of levels, results for aspects of language and mathematics tests will be reported as points on a scaled score, set with the average at 100. These scores indicate only where a child is in relation to others and nothing about where they are in progression in learning ... The new system for assessment and accountability for primary schools in England still suffers from over dependence on testing and the use of end of key stage two tests for too many purposes.' (Harlen, 2014, pp. 31-32).

UKLA is also concerned that there is insufficient understanding of the limits of the statistical models in use, which cannot sufficiently account for the varied paths to progress that children take. The current SATs are insufficiently grounded in a credible test architecture and have no clear connection to the mandated curriculum.

Accountability arrangements should be based on system of sampling of children to avoid the onerous and stressful testing experience that the current system

forces upon young children.

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.:

No. UKLA takes the view that teacher assessment should be reported at key stage 2, not just the results of the reading (and mathematics) tests. TA should draw on a wider range of evidence that has been collected to make possible a firmly evidence-based judgement. This is currently what happens in Y2, where the Y2 test is statutory, but constitutes only one (important) piece of evidence that is taken into account. Other relevant evidence may include written responses, guided reading notes, reading journals, assessment of fluency through running records etc. It is neither valid nor reliable for one test result from one day to be the only evidence on which a crucially important judgement is made.

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.:

YES. This is important for the reasons stated in answer to Question 15 above.

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.:

UKLA does not have a view on these matters.

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.:

UKLA does not have a view on these matters.

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.:

UKLA believes that while all pupils should be assessed for the important purposes of formative assessment and reporting to parents, assessment for the purposes of school, LA and national accountability need not and should not involve every pupil in the year group. It can and should be carried out through a process of sampling (Harlen 2014), since this permits a more thorough but less onerous operation, yielding richer information at less cost to those producing, collecting and processing it. Different children can sit tests of different aspects of their literacy learning, giving a fuller picture of the group's achievements. But, if the information it yields is to be of use in improving the effectiveness of the system, such testing also needs to be based on a rich, informed and valid view of what it is to be literate.

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.:

Yes.

We know from Harlen's extensive review of the literature (2005) that 'when deciding the method, or combination of methods, of assessment for summative assessment, the shortcomings of external examinations and national tests need to be borne in mind' (p.266). Harlen's work (2005, 2014) suggests that when schools develop an 'assessment culture' in which assessment is discussed constructively and positively, the most effective teacher assessment can occur. UKLA

believes that, as part of developing an 'assessment culture' for literacy, teachers, both within schools and across schools, need to be given the time to discuss children's progress in writing. They need to be clear about the learning goals and so moderate their judgments carefully and together in line with these goals. Professional development is crucial – to develop teachers' subject knowledge to improve assessment, pedagogy and learning. UKLA believes a clear framework for CPD needs to be developed that is properly funded and staffed. Local Authorities are best placed to provide this

Harlen (2005:263) concludes:

'It can be argued that for dependable summative assessment (i.e., with construct validity protected and optimum reliability) the requirements are:

- Decisions about the domain of knowledge, skills and other attributes of learning to be assessed that are justified in terms of how learning takes place.
- A valid sample of student behaviour in the domain.
- Criteria for judging the sample that are well matched to the goals of the work, of the curriculum and of the domain.
- Procedures for the reliable and unbiased application of the criteria.
- Procedures for reporting and communicating with users of the assessment outcomes.

UKLA supports the greater use of teachers' summative assessment for the assessment of writing, but with the proviso that teachers are encouraged and supported to meet such requirements through collaborative exercises carried out within and between schools, as outlined above.

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.:

As we have shown drawing on recent research, effective assessment for formative purposes provides information that is useful in three ways:

- for short-term decisions about learning and teaching,
- for reporting to parents (and discussion with pupils)
- for longer-term planning.

This is particularly true for writing.

Recently, in collaboration with UKLA and other associations and centres for the teaching of English, the Centre for Literacy in Primary Education (CLPE) has redeveloped the CLPE Reading and Writing Scales (<https://ukla.org/resources/details/clpe-reading-and-writing-scales>). These scales show multi-stranded progression, as children move from dependence to independence. They allow teachers to describe the complex and varied journeys that children take as they become more literate

The purpose of the scales is principally to help teachers to understand what progression looks like in reading and writing, to recognise what their pupils have achieved and to help them move forward. . At every stage, they include 'Next Steps' for the teacher and child to take. In this way the scales show what schools can do to support children's development as readers and writers, suggesting activities to take children into the next phase of their development.

The pedagogy underpinning the scales and the Next Steps is grounded in a coherent conception of children's language and literacy development, based on research evidence, which is clearly documented. Thus the scales provide a comprehensive framework for formative assessment in literacy, directly linked to effective forms of pedagogy. They also provide ways of documenting children's attainment and progress that are usefully informative for parents.

The two scales, one for reading and one for writing, each offer a description of observable behaviours at different stages. The writing scale documents a range of behaviours ranging from transcription of speech sounds and letter formation to awareness of the reader and the rehearsal and refinement of ideas. The reading scale is similarly comprehensive. Teachers will be able to think about where on the scales they could place the children that they teach. Once they have thought about this they can see the next set of observable behaviours that will show that the child is progressing in reading and writing. Using the scales to reflect on the attainment of children in their class gives teachers a clear idea about what to look for in day to day assessment and the key areas they need to plan for next. Every child takes a different journey through these scales. Their starting points and their rate and pattern of progression will depend on many factors including their prior experience, their interests and their learning preferences.

As to summative assessment, the higher the stakes the more necessary it becomes for such assessment to reflect the breadth of literacy and writing practices that have currency in our society and into which we are aiming to initiate our pupils. This means, inter alia, that any form of summative assessment must look beyond print on paper to broader forms of media in use in the wider society.

Lessons may be learned from the work of The National Educational Monitoring Project (NEMP) http://nemp.otago.ac.nz/_about.htm#1. Charged with examining New Zealand's progress in education within a literate society (Flockson and Crooks 1996), NEMP chose a sampling approach (as discussed at the answer to Q 15 above). . Its aim was to get a broad picture of the achievements of representative samples of New Zealand school students at successive points in time so that:

- trends in educational performance could be identified and reported;
- good information was available to assist policy makers, curriculum specialists and educators with their planning;
- the public could know about trends in educational achievement.

The assessment of literacy was a major part of this work. The team used a variety of innovative instruments to make their judgements., which included attention to writing at a computer, writing in students' own time, attitudes to writing and self-evaluation

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.:

UKLA agrees with the conclusions from Harlen (2014) that the current system for assessment and accountability for primary schools in England is over dependent on testing and the use of end of key stage two tests for too many purposes.

UKLA takes the view that teachers should be trusted to assess their pupils' achievement. This requires robust methods of moderation. There are a variety of ways of moderating teachers' judgements of children's attainment. High quality moderation requires teachers meeting together to discuss

examples of pupils' work. This way, teachers are able to share and discuss their interpretations of the performance descriptors and specific sets of children's works. Black et al (2013) report work with teachers aimed at developing teachers' assessment so that it can meet the requirements of reliability and validity. This approach is already a well-established form of professional development. This method also will enable trust and openness to be built up amongst all involved (Harlen 2014).

The UKLA believes strongly that information is needed from all areas of learning. This should also come from moderated teacher assessments which creates a more useful holistic picture of the child as a learner.

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.:

As Johnston and Costello, (2005:256) make clear, assessment is 'a social practice that involves noticing, representing, and responding to children's literate behaviours, rendering them meaningful for particular purposes and audiences'. It is important that assessment practices recognise that people engage in literate practices differently in different contexts. It must be recognised that the examination of the performance of a child on a particular task in a test situation is not always representative of all literate contexts. To assess a child's literate learning will require exploring not only what the child knows and can do but also the context within which the child can show they know and can do. If this perspective is not observed then the effect may be to seriously undermine a teacher's understanding of what children can know and do. There are many factors that may work to undermine the performance of a child in a particular assessment environment, for example, the extent to which materials being used are culturally relevant and accessible. It is important to shift the emphasis away from the isolated mind, to the mind in a social context (Johnson and Costello 2005).

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.:

UKLA would like to make a point about education policy and poverty. The effects of the latest deep education cuts and "reforms" to welfare have led to serious pressure on school and family budgets. These cuts have resulted in poverty becoming more widespread. We know from the Child Poverty Action Group that one in four children are growing up in poor households; the Institute for Fiscal Studies predicts a 50% increase in child poverty by 2021. Schools in the UK have tried to cover for the abandoned responsibilities of the state by offering food, clothing and stationary that many families are increasingly unable to afford as a consequence of the cuts.

In addition, The House of Commons Public Accounts Committee believes that the spending cuts jeopardise 'the quality of education and educational outcomes'. It notes also that the DfE has no means of monitoring their effect on educational standards. As Patrick Shannon (2014) has written about policy in the USA: 'If policy-making were a rational act, then...legislators would use the strong biological, psychological, and sociological evidence to design policies to provide material support for adequate food, shelter, income, and health care for all...(p.41)'. Poverty seriously limits children's capacity to take up the affordances of classroom opportunities to learn. Governments need to focus on ways to remove poverty as part of any aspect of education policy development in the UK which attempt to raise achievement in schools.

References

Black, P., Harrison, C., Hodgen, J., Marshall, B. and Serret, N. (2013) Inside the Black Box of Assessment. London: GL Assessment.

Bradbury, A. (2013) Understanding Early Years Inequality: Policy, Assessment and Young Children's Identities. London: Routledge

Bradbury and Roberts-Holmes (2016) "They are children...not robots, not machines": The Introduction of Reception Base line Assessment. London: National Unions of Teachers

Clark, M. M. (2015) A Critique of Synthetic Phonics...this is the evidence. Primary First pp. 22-26.

DfE (Department for Education). (2010) The Schools White Paper: The Importance of Teaching. London: DfE.

DfE (Department for Education). (2011) Statutory Framework for the Early Years Foundation Stage. Accessed July 10, 2013. <http://www.education.gov.uk/consultations/index.cfm>

DfE (Department for Education). (2012) EYFSP Attainment by Pupil Characteristics in England: Academic Year 2011 to 2012. Accessed May 10, 2013. <https://www.gov.uk/government/publications/eyfsp-attainment-by-pupil-characteristics-in-england-academic-year2011-to-2012>

DfE (Department for Education). (2013) More Great Childcare. Raising Quality and Giving Parents More Choice. Accessed May 5, 2013. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/219660/More_20Great_20Childcare_20v2.pdf 314 G. Roberts-Holmes

DfE (Department for Education). (2014) Reforming Assessment and Accountability for Primary Schools. Accessed March 10, 2014. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/297595/Primary_Accountability_and_Assessment_Consultation_Response.pdf

DfES (Department for Education and Skills). (2007) Early Years Foundation Stage: Setting the Standards for Learning, Development and Care for Children from

Birth to Five. Nottingham: DfES.

Elliott, V., Baird, J., Hopfenbeck, T.N., Ingram, J., Thompson, I., Usher, N., Zantout, M. with Richardson, J. and Coleman, R. (2016) A marked improvement? A review of the evidence on written marking. University of Oxford Department of Education and Education Endowment Foundation. Available on: https://educationendowmentfoundation.org.uk/public/files/Publications/EEF_Marking_Review_April_2016.pdf Accessed 8.10. 2016

Flockson, L. and Crooks, T. (1996) National Education Monitoring Project: Reading and Speaking: Assessment Results: 1996 (no.6) Dunedin, New Zealand: Educational Assessment Research Unit

Harlen, W. (2005) Trusting teachers' judgement: research evidence of the reliability and validity of teachers' assessment used for summative purposes, *Research Papers in Education*, 20:3, 245-270, DOI: 10.1080/02671520500193744

Harlen, W. (2014) *Assessment, Standards and Quality in Learning in Primary Education: A Report for the Cambridge Primary Review Trust*. Cambridge: Cambridge Primary Review Trust.

Johnson, P. and Costello, P. (2005) *Principles for Literacy Assessment* *Reading Research Quarterly*, 40: 2, pp. 256-67

Lingard, B., W. Martino, and G. Rezai-Rashti. 2013. "Testing Regimes, Accountabilities and Education Policy: Commensurate Global and National Developments." *Journal of Education Policy* 28 (5): 539–556

Myhill, D, Jones, S.M., Lines, H and Watson, A.M. (2012) Re-thinking grammar: The impact of embedded grammar teaching on students' writing and students' metalinguistic understanding. *Research Papers in Education*, 27:2, pp. 139-166

Roberts-Holmes, G. (2015) The 'datafication' of early years pedagogy: 'if the teaching is good, the data should be good and if there's bad teaching, there is bad data', *Journal of Education Policy*, 30:3, pp. 302-315, DOI: 10.1080/02680939.2014.924561

Shannon, P. (2014) 'Re-Reading Poverty; Reorienting Educational Policy' in K.S.

Goodman, C.R. Calfee, M.Y. Goodman (eds.) (2014) *Whose Knowledge Counts in Government Literacy Policies? Why Expertise Matters*. Abingdon: Routledge

Walker, M., Bartlett, S., Betts, H., Sainsbury, M. & Worth, J. (2015) *Phonics Screening Check Evaluation: Final report*. Slough: National Foundation for Educational Research. Accessed 27.10.2016 at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/434821/RR418A_Phonics_screening_check_evaluation.pdf