

PRIMARY ASSESSMENT IN ENGLAND GOVERNMENT CONSULTATION MARCH 2017

The Australian Council for Educational Research (ACER) is pleased to submit this response to the Department for Education (DfE) and the Standards and Testing Agency (STA)'s open consultation, *Primary assessment in England*.

From Professor Geoff Masters, AO, Chief Executive, ACER Ltd and Martin Ripley, Non-executive Director, ACER UK

1. THE AUSTRALIAN COUNCIL FOR EDUCATIONAL RESEARCH (ACER)

ACER is a not-for-profit organisation, set up in 1930 with 87 years of experience in assessment around the world.

In the UK, ACER UK is developing and running the Scottish National Standardised Assessment (SNSA) programme for the Scottish government. This is an innovative assessment programme with several commendable features to which we return later in this response.

ACER has extensive expertise in national and international assessments, having managed the global PISA programme from its inception until 2012. In addition, we have made contributions to the International Association for the Evaluation of Educational Achievement (IEA)'s *Trends in International Mathematics and Science Study* (TIMSS) and to UNESCO's global education indicators programme.

Using this background and accumulated expertise in response to the consultation, ACER draws on some of the major findings and conclusions of our experience of national assessment programmes. These findings and conclusions include the effects that system-level approaches to assessment can have on the progress made by pupils and the ways in which schools adapt their practices and behaviours in response to national assessment approaches.

2. PURPOSES

Primary assessments in the Early Years Foundation Stage Profile (EYFSP), Key Stage 1 and Key Stage 2 are designed to achieve multiple purposes: school accountability, measuring pupil progress and evaluating individual children's learning needs. In our experience, the use of assessment to monitor learning can be contrasted with more traditional uses of assessment to judge and grade pupils.

These two approaches to assessment differ in several important respects.

First, they differ in *intention*. The intention in grading is to use children's assessment performance to reach judgements about how well they have learned what they have been taught. The onus is on pupils to learn what teachers teach, and assessment is the process by which teachers are held to account. In contrast, the intention in monitoring is to gather information that can be used to promote further learning and improved learning outcomes. Monitoring is viewed as an integral part of professional work and effective educational decision-making.

Second, these two approaches differ in *focus*. The focus in grading is on the content of a particular area of the curriculum (e.g. a topic, learning module or school year). Assessments are undertaken to establish how well pupils have learned the prescribed curriculum content and whether they have reached pre-determined expectations (or standards). In contrast, the focus in monitoring is on pupils' long-term progress in a domain of learning. Assessments are undertaken to establish what pupils know, understand and can do at the time of assessment.

Third, these two approaches differ in the *priority* they assign to *year-specific expectations*. In grading, pupils who perform well on year-specific expectations receive high grades; those who perform poorly receive low grades. All pupils are held accountable for achieving common end points despite the fact that they might have very different starting points. In contrast, in monitoring the priority is to establish pupils' levels of achievement in a domain regardless of their year level—that is, in an absolute sense. This does not preclude also setting expectations of pupil achievement by particular times in their schooling.

Fourth, and as a result, the two approaches differ in their *interpretation* of what it means to *learn successfully*. In grading, successful learning is defined in terms of year-specific curriculum expectations and usually is summarised in a level or a score. A pupil can receive the same level or grade (e.g. 'emerging', 'expecting' and 'exceeding' in the EYFSP). Pupils can find that they are below expectation at every assessment. Other pupils may achieve above expectation, but receive no information about the progress they have made or whether they are 'coasting' despite achieving well. Such feedback disguises the progress children actually make over time and sends a message about their ability to learn (they are an 'underperforming' pupil). In contrast, in monitoring, successful learning is defined as progress or growth towards greater knowledge, deeper understanding and more advanced skills, regardless of starting point.

3. LEARNING DOMAINS

The process of establishing where learners are in their learning depends on a deep understanding of the relevant learning domain.

An analogy would be the monitoring of a person's progress through physical terrain. An appreciation of where an individual is in their progress depends on familiarity with, and at the least having a mental map of, the terrain through which they are travelling. Understanding a learning domain includes understanding typical sequences and paths of learning; appreciating how learning builds on earlier learning and lays the foundations for further learning; understanding the importance and role of prerequisite skills and knowledge to learning progress; and familiarity with common pupil errors, misconceptions and obstacles to learning. Such understandings— sometimes referred to as pedagogical content knowledge—are developed through professional experience and informed by domain-specific learning research.

Progress in an area of learning usually takes the form of increasing *proficiency* reflected in more extensive knowledge, deeper understandings and higher-level skills. Establishing where pupils are in their learning at any given time means clarifying what they know, understand and can do at that point in their progress. This can be done at the level of an entire learning domain such as mathematics, for particular sub-domains such as number, space/geometry and algebra, or in relation to specific skills and understandings.

Most learning domains extend over many years of school. Some begin before, and continue beyond, school. Examples include the broad domains of reading, mathematics and science. Most general attributes and capabilities, such as the ability to work in teams, to communicate, and to create innovative solutions to problems, also develop over extended periods. Because children of the same age tend to be at widely varying stages in their learning progress, teachers require a good understanding of the nature of long-term development if they are to ensure that every pupil is appropriately challenged and extended.

An example of a learning domain for which increasing proficiency has been well defined is music. This domain often is divided into sub-domains such as theory of music, music craft and musicianship, with pupil progress being monitored and recognised in each sub-domain separately. Within each sub-domain, increasing proficiency is defined by a series of levels or 'grades'. Pupils progress through these grades at their own pace, demonstrating that they have achieved each grade at a time of their own choosing—a process that occurs over a number of years. When pupils understand what higher levels of proficiency and better quality work look like in an area of learning, they have a basis for reflecting on their current levels of achievement and for setting realistic, challenging goals for further learning. In this way, children are supported to take a degree of responsibility for their own learning and long-term progress.

It is common to interpret progress through a learning domain by reference to the points pupils are expected to reach by particular stages of schooling. For example, the minimum

level of reading proficiency expected of all pupils by the end of Year 3 might be specified. The specification of minimum expectations (or 'standards') leads to *yes/no* decisions about whether pupils have achieved these expectations. In education systems around the world, policy and performance targets have been built around such *yes/no* thresholds, with mixed results.

4. TECHNOLOGY

Your consultation did not ask specifically about technology. However, you asked about reducing burden and workload in foundation stage (Q3); about burdens and workload in relation to the multiplication tables assessment (Q14); and about administration of statutory assessment (Q15).

We noted that your consultation did not ask about ways in which assessments could be made more useful to learning.

We believe that further consideration should be given to the potential role of technology in the design, delivery and reporting of assessments. We note that the national assessments being introduced in Scotland are computer based, and that early feedback is suggesting that headteachers and schools are delighted with the ease of administration, the quality of the test content in literacy and numeracy, and most especially with the detailed and diagnostic reporting available for schools, class teachers and individual learners.

Online testing with immediate results and reporting will require the least work by the school. For multiplication tables assessment, which is a well-defined content area, the test can be fairly short, even as little as 5 minutes. We would suggest that for each pupil the test selects questions randomly from a question bank. This would reduce the possibility of 'cramming', which is known to increase workload, and enables reliable testing even if the school does not have the facilities to organise formal examination conditions online, but pupils must work close to each other.

Online assessment will also enable formative feedback, for example highlighting the times tables the pupil does not master yet, or those that they know very well, which could be identified based on response time.

Digital assessment reduces the administrative burden of assessment significantly. In addition, it provides possibilities for more informative feedback, which can help the teacher and can be used in the classroom. This would add benefits to schools and teachers in reducing the perceived inconvenience of the statutory assessment process. Digital solutions also often allow better accommodations to be made for learners with special needs.

Assessment systems such as the Scottish National Standardised Assessment (SNSA) programme offer compelling evidence of the capacity of primary schools to deliver computer-based assessments to all pupils. We believe that the DfE should consider

suitable 'exploratory' approaches. For example, the multiplication tables assessment could be designed and delivered to schools as a highly flexible assessment, designed to monitor and measure progress, helping schools to identify those pupils not making progress and to support them effectively.

5. TIMING OF ASSESSMENTS

Teachers require good understanding of where children are in their learning (current levels of knowledge, skill and understanding) to design learning opportunities appropriate to individuals' and groups' levels of readiness and learning needs.

Evidence from educational assessment programmes designed to establish where pupils are in their long-term learning is revealing enormous variability in learners' levels of achievement. Research suggests that this variability is due in part to individuals' past experiences and to the environments they have been exposed. Learning, then, is a highly personal process, with the consequence that individuals of the same age can be at very different points in their learning and development. By the time children commence school, they have markedly different levels of cognitive, language, emotional, social and psychomotor development.

It is therefore crucial that children's levels of development and learning are accurately identified at the start of school, and that the progress made due to the experiences and the environment provided by the school is recognised. This not only supports the development and learning of individual children and schools, but also in the longer term allows for effective practices to be identified and applied throughout the education system.

In our view, the key characteristics of baseline assessment should be as follows.

Baseline assessment should be *balanced*:

- covering a sufficiently wide range of skills and knowledge for all children to be able to show what they know, understand and can do;
- covering a sufficiently wide range of skills and knowledge for the acquired assessment information to be useful, but not a burden on children or teachers.

Baseline assessment should be *fair*:

- taking into account the age of the children in terms of assessment and administration method, and content;
- taking into account the individual development of each child;
- used to inform, not to judge.

Baseline assessment should be *uncomplicated*:

- easy to administer;
- capable of being reliably administered by the class teacher and/or other teaching staff;

- with results that are easy to understand;
- with results that inform what teachers do in the classroom, and that are easy to apply without excessive work.

We would also suggest that the reception baseline assessment should complement the EYFS framework to avoid unnecessary duplication of work, and to ensure that the combined assessment data provide a good breadth of information.

Our considered response to each question in the open consultation follows.

ANNEX 1 – DETAILED CONSULTATION RESPONSE

Q1. 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.

We believe that the ELGs set out in the EYFS framework are robust, but that there remains considerable doubt about how these map onto learning pathways in key stage 1 and key stage 2. This would help address, for example, how best to support a Year One child who has performed well in a phonics-based reading learning programme but who has difficulty in breaking through to establish comprehension of what has been read.

This response is informed by the observations above in Section 3.

Q2. 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?

No, we don’t believe that categorising children into ‘emerging’, ‘expecting’ or ‘exceeding’ is as helpful to learning as measuring and describing the progress they make over time.

This response is informed by the observations above in Section 2.

Q3. What steps could we take to reduce the workload and time burden on those involved in administering the EYFSP?

Although there will be a number of areas to consider in reducing workload, we would particularly want to emphasise the points we have made above. EYFSP data is only useful to learners to the extent that it relates to agreed learning progressions – domain and sub-domains that clearly map into the KS1 curriculum and where children’s progression in learning can continue to be mapped. Other assessment data provide limited value to any learner.

Second, we believe that a further and more sustained examination of the efficiencies benefits of technology should be considered – in designing and delivering assessment, and providing results data. Third, we believe significant further strides could be taken by re-thinking the reporting functions of EYFSP – reporting to teachers and parents in terms of progress made by children and information that is ‘classroom-ready’ in a form and at a time that will be helpful to teachers.

This response is informed by the observations above in Section 4.

Q4. How could we improve the consistency and effectiveness of the EYFSP moderation process whilst reducing burdens?

We have no specific response to this question.

Q5. 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.

Provided the assessment domain is clearly defined and that progression within that domain is described and widely understood, we agree that the first assessment can occur within reception. The problem with this, as we note above, is that learning progressions from reception domains into key stage 1 assessment subjects has not been described – it is not clear how the entirety of EYFSP maps into the existing key stage 1 assessments.

This response is informed by the observations above in Section 5.

Q6. 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.

We agree with the government's suggestion of administering the assessment at the beginning of the second half term. This will give children some time to grow comfortable with their new environment, which allows them to do their best, while also being early enough in the year to provide a reliable measuring point for overall progress in primary education.

This response is informed by the observations above in Section 5.

Q7. 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?

We agree.

Q8. 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?

We agree.

Q9. 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?

No, we believe that at least one statutory assessment is required before Year 6.

Q10. 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?

International evidence indicates that school systems that use pupil performance outcomes as the sole, or main, measure of accountability have either regressed or made markedly less progress than systems that combine measures of pupil performance with other school level measures, including quality of teaching, CPD and investment in staff.

We strongly recommend that consideration should be given to unintended consequences of accountability measures based on pupil performance in England and to international evidence indicating that a wider and more balanced set of measures promotes the longer term health of education systems, including by fostering sustainable improvements in pupil performance.

Q11. 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. We think teacher assessment should continue, although an alternative option would be to make it 'recommended' rather than statutory.

Q12. 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. The current freedoms and flexibilities enable teachers to choose the approach that best meets the needs of their pupils.

Q13. At what point in key stage 2 do you think the multiplication tables check should be administered? Please explain the basis for your views.

- a) At the end of year 4
- b) During year 5
- c) During year 6.

Our recommendation is for *Option B*, during the first term of year 5.

This would ease the end-of-year burden in schools and ensure that children still remember their times tables after the summer break, that is, that learning has become embedded. It would also allow schools to immediately provide support and interventions to pupils who require it.

This response is informed by the observations above in Section 5.

Q14. How can we ensure that the multiplication tables check is implemented in a way that balances burdens on schools with benefit to pupils?

We believe that a further and more sustained examination of the efficiencies benefits of technology should be considered.

This response is informed by the observations above in Section 4.

Q15. 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?

We believe that a further and more sustained examination of the efficiencies benefits of technology should be considered.

This response is informed by the observations above in Section 4.

Q16. 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.

No comments.

Q17. Please give details of any robust alternative approaches to the assessment of English writing, which the Department for Education should explore.

No comments.

Q18. Please give details of any effective models of moderation or standardisation of teacher assessment that the Department for Education should explore.

We recommend that you review the Scottish National Standardised Assessment programme.

Q19. Do you think that any of our proposals could have a disproportionate impact, positive or negative, on specific pupils, in particular those with 'relevant protected characteristics' (including disability, gender, race and religion or belief)? Please provide evidence to support your response.

No comments.

Q20. 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.

No comments.