



**Evidence for
Excellence in
Education**

Consultation response

Secondary school accountability consultation

May 2013



About NFER

NFER is the UK's largest independent provider of research, assessment and information services for education, training and children's services. We make a difference to learners of all ages, especially to the lives of children and young people, by ensuring our work improves the practice and understanding of those who work with and for learners.

NFER's purpose is to provide independent evidence which improves education and training. As a charity we are always thinking about ways in which we can achieve this purpose. In particular, we are focusing attention on the use of evidence for improvement, taking a proactive approach to setting the research agenda for education and children's services in order to make a real impact on policy and practice.

About this document

Between February and May 2013 the Department for Education held a consultation on secondary school accountability. This document is NFER's response.

The paragraph numbering in the document follows the consultation questions. We only answered those questions where we could draw on recent work that provided evidence or expertise.

The consultation set out the aims and vision for the proposed new accountability system, together with the case for change. The proposals reflect the changes to GCSE exams, and focus on issues around the publication and use of data.

Related documents

Other NFER consultation responses and policy papers relevant to the development of secondary education are also published on the NFER website.

NFER
May 2013
www.nfer.ac.uk

1. Do you agree with the proposals for the headline accountability measures?

It would be valuable to have an agreed purpose for secondary schools in England as this would then make the questions about how the schools should be held to account much easier to answer. Do the schools exist to develop the academic skills of their pupils, or do they have a much broader remit in terms of developing well-rounded individuals? Are we aiming to prepare pupils to be responsible citizens, effective parents, and conscientious employees, or are we aiming to teach them an agreed set of knowledge and understanding? Countries such as Singapore have very clearly defined purposes for their education system, and then accountability and policy developments are lined up to achieve those purposes. In the case of Singapore, this is expressed in terms of "[Desired Outcomes of Education](#)".

As we do not have this agreed set of purposes, the following responses, by their nature, reflect an inherent belief about what our secondary schools should deliver. NFER is planning to conduct some work around the purposes of education, but we are supportive of the World Bank view that 'education enhances people's ability to make informed decisions, be better parents, sustain a livelihood, adopt new technologies, cope with shocks, and be responsible citizens and effective stewards of the natural environment' ([2020 Education Strategy](#), World Bank). We would be supportive of a set of purposes aligned towards these broad aims.

Given the above assumptions:

We are supportive of the proposal to hold schools to account for the development of solid foundations in literacy and numeracy in all pupils.

Similarly, we believe that a focus on the best eight subjects for pupils will allow them to engage with a broad curriculum, across a range of subject areas, while not encouraging schools to push pupils to attempt a large number of subjects at the expense of quality in the core. An average point score 8 measure is a good idea in this context because it is a continuous measure rather than a borderline measure, so should encourage schools to move away from any current focus on pupils at the C/D borderline. The average point score 8 measure will also provide recognition of achievement for those pupils who may not achieve a grade C or above, such as those with special educational needs for whom, in some cases, a grade D or E may be a real achievement.

We agree that the eight subjects should include a particular focus on literacy and numeracy, and then a broader range of good quality qualifications. We also agree that a broader range of subjects, both general and vocational qualifications, should be recognised in the eight.

We agree that the introduction of a sample test at key stage 4, to monitor standards of the education system over time, would also be valuable. This should be done in a way that is of the lowest possible burden to schools, for example it is not likely to be necessary to administer the test annually – for changes in standards a 3-year cycle is possible. Similarly, there are different ways in which the sample can be selected, and we would argue for a lower number of schools with a larger number of pupils than what we understand to be

currently proposed for the key stage 2 sample tests. The increase in reliability afforded by the current sample method is unlikely to be sufficient justification for the increased burden on schools and the increased cost of the model.

We are not convinced by the need for the EBacc measure, nor of the current list of subjects included in the measure. We believe it currently includes too limited a range of subjects and will encourage schools to narrow the curriculum offered to pupils, such as by limiting the range of art subjects on offer. We do not think this is useful in providing a curriculum that best suits the needs of all pupils or that develops the 'well-rounded' citizens that ought to be the result of our education system. See: [Key stage 4 policy reform and its potential impact on at-risk young people](#).

Part of the proposal is to publish 'extensive data' about secondary schools through the data portal. We support making data available and being transparent about what schools deliver in a range of areas. However, by making a huge amount of data available it can sometimes make it less useful, rather than more useful. If there is a huge amount of data it may mean that only experts are able to find what is really useful among the detail. By including a focus on a number of key measures, it is likely to be that little use is made of the additional data. It would be useful to monitor, over time, what data is used and by whom, to inform the data that continues to be made available. Additional data can still be available for research or other purposes.

2. Is there any further information we should provide about the performance of disadvantaged students?

We think that having a broad curriculum recognised through the eight subjects, with the option to include a range of good quality general and vocational qualifications as part of this, should ensure that schools are encouraged to provide a broad and balanced curriculum suitable for the majority of pupils, including those from disadvantaged backgrounds. The priority of the education system should be that all pupils achieve their full potential, in a way that gives them access to the best next steps, and ensures that they make as much progress as possible.

However, it is possible to disguise poor performance by one group with good performance by another in an average measure, so there should be an explicit way of looking at narrowing the gap success. This could be by providing evidence about the performance of different disadvantaged groups or perhaps, for example, by publishing the difference in average progress between the bottom 20% and the top 20% in the school.

3. Should we look to use a relative measure as the floor standard in the first year of the new exams?

From the proposed timescales of the various changes underway in schools, it is clear that the pupils studying the new GCSEs will have followed the same curriculum for the majority of their school life as the pupils taking the current GCSEs. There may be some changes in attainment possible due to taking fewer modules during a course, but there is little reason to

assume that they will be able to achieve a very different standard just because they are taking a harder exam. For this reason, it seems likely that pupils at the lower end of the C grade are not likely to achieve the standard for the C grade in the new GCSE. In addition, the fact that the style of the examination is changing is also likely to mean that achievement will fall until teachers and pupils become familiar with the new format, and a gradual improvement will be seen over the first few years after its introduction.

For these reasons, it is likely that the new GCSEs will have a negative impact on thousands of young people in the first years after its introduction. In order for all pupils to have equal chance of achieving a grade C (a very high stakes benchmark) over time, it is necessary to award the new qualifications at the same standard (possibly using the comparable outcomes approach) in the short term. A decision about how and when the new standard can be applied should be considered once evidence of achievement is available.

4. Are there any other measures we should consider publishing?

We think it may be useful to report some additional measures (see response to question 8) depending on the shared view of the purposes of secondary schooling. The reported measures are given a particular focus in planning the work of a school, and as such if we believe that additional skills, such as developing responsible citizens, are important, then it would be essential to have measures that reflect such skills. It would also be valuable to publish measures such as parent, pupil and teacher satisfaction.

In 2009, NFER published a paper on the information published about schools in other countries and made recommendations about what ought to be published about schools in England. The findings from this paper largely continue to hold true. The paper is available [here](#).

5. Do you think we should collect and publish test data from internal assessments through the Data Warehouse?

See the comments above about too much data being available. There are also likely to be some issues with the quality of the data from internal assessments that mean it is not comparable to the data produced externally. This may not be well understood by the users of the data, which may impact on conclusions drawn from it. The publishing of internal assessment data in this way may also encourage schools to focus on summative assessment rather than on formative assessment. There is significant evidence that formative assessment can have a positive impact on pupil attainment (Black and Wiliam, 1998).

7. Do you agree that the Department should stop the collection of Key Stage 3 teacher assessment results?

Key stage 3 teacher assessment results are not required for the current accountability proposals and it is not clear what value they would add were they to be reported as part of

the measures. We are not aware of research into the reliability of the results, nor into their perceived usefulness by stakeholders. Without this evidence, it is difficult to say whether they hold particular value to certain groups, or if they are used for important reasons by certain stakeholders. We are not aware of any reason to continue with the data collection.

8. How should we ensure that achievement beyond formal qualifications is recognised?

Given the assumed broader purposes of secondary schools, as described above, we do believe it is essential that achievement beyond formal qualifications is recognised. This is important in that it will give a value to the additional learning for those who are planning the school curriculum, for young people and their families who are deciding on their areas of learning, and for those who will work with the young people in the next stages, such as FE, HE and employment.

There are a number of possible ways that achievement can be recognised beyond the formal qualifications, such as some well-established systems of 'badges', that recognise achievement in a range of different areas. The Duke of Edinburgh Award is a good example of this. These badges, or similar approaches, can be used to build up a personal 'statement of achievement' for all pupils. New technology could be used to facilitate this process. The focus of this consultation is accountability of schools, and as such badges at an individual level may not add significant value. However, some ways of recognising the broader contribution of the school are likely to be of value (if our views about the purposes of secondary education are shared). This may be part of Ofsted's remit.

We believe that Ofsted has an important role in evaluating how well schools develop a full range of skills in their pupils and how they ensure that all pupils develop to their full potential across a range of areas, including academic, but also including a much broader range of skills, which may include: leadership, citizenship and community engagement. Ofsted should also consider how schools perform in areas such as pupil and teacher well-being and parental satisfaction.

9. How can national sample tests best be introduced?

In 2009, NFER published the findings from a seminar on the issues to consider when introducing a national sample test. The findings can be downloaded [here](#) (NFER, 2009).

As mentioned above, we think the national sample tests should be introduced in such a way that they have minimum impact on schools, while still being conducted in a reliable way. This means that:

- The sample should be designed to be as reliable as required while not placing a burden on too large a number of schools;
- The frequency of the surveys should be appropriate for the size of change likely to be measured, e.g. it is unlikely that annual surveys will be required as changes in standards year by year are likely to be small;

- The number of subjects included should be broad enough to reflect what is considered to be important (while recognising that there is likely to be a high correlation between similar subjects) but narrow enough not to over-burden schools. It is likely that sample tests in a small number of subjects could be linked to GCSE results in such a way that a large number of subjects are not needed;
- The tests themselves should be designed to reflect the skills and understanding defined in the national curriculum, and in the GCSE specifications, e.g. if practical skills are considered to be important in a subject, then they should be assessed in the sample tests;
- The tests should be developed to cover the whole curriculum, by using a large number of question papers and a 'matrix design' when allocating the question papers to different pupils.

NFER would be happy to be involved in discussions about the design of the new sample tests and could bring valuable expertise from our experience of the Assessment of Pupil Performance sample tests and from our involvement in the international surveys such as PISA, TIMSS, PIRLS and the European Languages Survey.

References

Black, P. and Wiliam, D. (1998). 'Assessment and classroom learning', *Assessment in Education*, **5**, 1, 7–74.

National Foundation for Educational Research (2009). *Submission to Expert Group: Issues to Consider when Developing a National Monitoring System*. Slough: NFER [online].

Available:

http://www.nfer.ac.uk/nfer/publications/99901/99901_home.cfm?publicationID=330&title=Submission%20to%20expert%20group:%20issues%20to%20consider%20when%20developing%20a%20national%20monitoring%20system [3 May, 2013].

NFER provides evidence for excellence through its independence and insights, the breadth of its work, its connections, and a focus on outcomes.

- independent
- insights
- breadth
- connections
- outcomes

**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