

Technical and methodological statement for key stage 2 science sampling

February 2010

QCDA/10/4773

QCDA is working in partnership with the Department for Children Schools and Families and this document has been jointly prepared by statisticians.

Introduction

In May 2009 the Government accepted in full the recommendations of the Expert Group on Assessment, a group comprising headteachers and assessment experts.

The group recommended that, while key stage 2 science tests had played an important role in supporting improvements in teaching and learning, they should now be discontinued. From 2010, key stage 2 science at school level will therefore be assessed by teacher assessment only, which takes greater account of pupils' practical grasp of the subject and is based on their attainment throughout the academic year across the full programme of study. As a result key stage 2 national curriculum science tests have been discontinued. From 2010, key stage 2 science at school level will be assessed by teacher assessment only.

In order for the Department for Children, Schools and Families (DCSF) to monitor national standards, there will be externally-marked science sampling tests, administered by the Qualifications and Curriculum Development Agency (QCDA). DCSF and QCDA will not report on the attainment of individual local authorities, schools or pupils.

The purpose of the methodology statement

This statement has been prepared jointly by statisticians at the QCDA and DCSF. It explains how schools have been selected to participate in the science sampling assessment.

The statement contains the technical and methodological details of the key stage 2 science sampling exercise. Specifically, the statement provides detail about how the sample was drawn and how results will be reported. The statement also identifies issues that should be considered when interpreting results.

This statement is written for the academic community, such as statisticians and psychometricians with an interest in sampling arrangements to monitor pupil attainment.

Selected schools and their local authorities may also find the statement of interest, to further explain how the sample was selected. Further information for schools and local authorities about science sampling and its administration is available on the QCDA website at www.qcda.gov.uk/sciencesampling.

The sampling exercise

The sampling exercise has been designed to enable the identification of the proportion of pupils in the population who would attain level 4 or above in the 2010 key stage 2 science tests, on the basis of the performance of a sample. The assessment instrument to be used will be the 2010 key stage 2 science tests, which have been developed in accordance with the requirements of the Ofqual code of practice for full cohort national curriculum assessments.

Reporting

National level reporting: DCSF will report results at national level only close to the beginning of the autumn term. DCSF intends to report the overall proportion of pupils attaining level 4 or above, and by gender. DCSF also plans to report the overall proportion of pupils attaining level 5, and by gender.

Local authority level reporting: no results will be reported at local authority level as the sample has been designed to produce robust results at national level only.

School level reporting: an individual school will receive feedback on the performance of its pupils, in terms of the average performance of the pupils, at attainment target level, compared with the average performance of the aggregate of all pupils participating in the sample, again at attainment target level.

This sampling exercise will not be used for school/local authority accountability purposes.

Description of the sample

The sample has been designed to enable comparison of an estimate made of the proportion of pupils attaining level 4 or above in the 2010 science sampling tests against the equivalent proportion of students attaining level 4 or above in the whole cohort tests administered in May 2009.

As proportions are published to the nearest percentage point the sample has been designed such that a difference in the above of 1 to 1.5 percentage points can be identified as statistically significant.

The approach is intended to avoid imposing disproportionate administrative burden onto the sector. An upper ceiling of 5% of schools was imposed to reduce both the number of schools involved in any one year, and the probability of a school being selected numerous times in a small number of years.

Historic data has been used to estimate, over repeated sampling, the standard deviation necessary for a 1 percentage point change to be significant at the 95% confidence level. In this instance the standard deviation should be no more than 0.36 percentage points.

For operational reasons the tests will be administered to the complete cohort within a school who have reached the end of the key stage 2 programme of study, which creates a clustered sample design and an associated increase in sampling error.

The selection of a simple random sample of schools of, on average, 632 with around 23,750 pupils, would give a standard deviation of 0.46 percentage points, which is in the mid-point of the above requirement (of 1 to 1.5 percentage points). To check and control for bias it is helpful to analyse whether the level 4 or above proportion correlates with other factors. It has been found not to correlate with the number of schools, pupils or the gender split in the sample. However, there is a correlation with the proportion of pupils eligible for Free School Meals (FSM). This is a known strong effect that it is desirable to control for; therefore, the sample will be stratified into 5 FSM bands by sorting schools in order of proportion of FSM eligibility and cutting this list into 5 approximately equally sized strata from which an equal number of sample schools will be randomly drawn. We estimate this could allow a 1.1 percentage point change in the level 4 and above proportion to be identified as statistically significant. The equivalent standard deviation in the level 5 proportion is approximately twice as large; therefore, for the level 5 proportion, only changes in excess of 2 percentage points can be identified as statistically significant.

For operational reasons a contingency (c20 per cent) has been added to the estimated sample size to ensure that factors such as schools closing before the end of the academic year, or losses of scripts in the postal system do not put the ability to produce an estimate with the desired level of accuracy at risk. Therefore, a sample of 750 schools has been selected, which is expected to yield a sample size of approximately 27,099 pupils.

The selection of the sample

The population of schools that were sampled was drawn from the 2009 Autumn School Census and consisted of all maintained schools that returned at least one pupil in National Curriculum year 6 (i.e. pupils that are typically at the end of the key stage 2 programme of study) in England. Independent schools are not included in the sample. Information on the proportion of pupils eligible for Free School Meals within each school has also been drawn from the same school census.

DCSF also requested that schools that are participating in the Single Level Test extended pilot be excluded from the population of schools. This is to ensure that no disproportionate burden is placed upon them through invitations to participate in evaluation of both the Single Level Test extended pilot and the science sampling exercise. As these schools represent a very small proportion of the number of schools nationally, this is unlikely to introduce any significant bias into the final estimate of the outcomes.

Issues around the interpretation of the analysis of the outcomes from the key stage 2 sampling tests in 2010

The sampling exercise has been designed to enable the identification of the proportion of pupils in the population who would attain level 4 or above in the 2010 key stage 2 science tests, on the basis of the performance of a sample. The assessment instrument to be used will be the 2010 key stage 2 science tests, which have been developed in accordance with the requirements of the Ofqual code of practice for full cohort national curriculum assessments.

In 2010 pupil and school level results for key stage 2 science will be based on teacher assessment judgments only, rather than on the results of a national curriculum test. Schools planning and teaching of key stage 2 science may have been modified to reflect the new arrangements for assessment at the end of key stage 2. Therefore, this sampling exercise will identify the proportion of pupils in the population who would have attained Level 4 or above in the 2010 key stage 2 science tests had they been administered to all pupils in maintained schools for the purposes of monitoring national standards, but not of accountability.

The quality assurance of the process

Quality Assurance (QA) has been a fundamental part of drawing the sample. Statistical quality assurance has taken place to ensure that the sample has been drawn appropriately and meets the DCSF's requirements. The process quality assurance put into place ensures that the sample of schools selected is managed appropriately within QCDA. Specifically, the quality assurance arrangements enacted ensure that:

The school census data file passed to QCDA was up-to-date and complete;

- The sample selection requirements were clear and unambiguous;
- Sufficient detail about eligibility for Free School Meals was given, to enable QCDA to stratify the sample;

- The sample selection process and the final sample were validated; and
- The sample selection met DCSF's requirements.

Additional QA will occur following the sample selection. Statistical and process QA will be applied to ensure that:

- The correct level thresholds are applied to the results file;
- The tests results files are complete and accurate; and
- Accurate feedback at individual school level is provided.

Ofqual's role is to monitor all aspects of the assessment process, from the development of tests to the collection and reporting of results data. Their code of practice sets the criteria for this.

Relevant links:

National curriculum assessment: www.qcda.gov.uk/tests

Science sampling: www.qcda.gov.uk/sciencesampling

Ofqual: code of practice for national curriculum assessments