The National Reference Test in 2017

Introduction

Each year, the National Reference Test (NRT) is taken by a sample of students in year 11, shortly before they take their GCSEs. This year's test was held in 341 schools between 20 February and 3 March. At each school, up to 30 students took the English test and another 30 students took the maths test.

Ofqual introduced the National Reference Test (NRT) to provide information on changes in performance over time in GCSE English language and maths. It is based on results from a nationally representative sample of students who take the test shortly before taking their GCSEs. From 2019 onwards, the information from the test will be considered each year by Ofqual and the exam boards prior to GCSEs in English language and maths being awarded.

In September, we published an Overview[1] of the NRT including information about its design and development, and how it is delivered each year. Ofqual awarded a contract to the National Foundation for Educational Research (NFER) in 2015 to support our introduction of the test and to deliver it on Ofqual's behalf. NFER carries out the testing in each school, and marks and analyses the students' papers. At Ofqual's request, in September 2016 the Government introduced legislation that makes it mandatory for most schools to take part in the test if they are included in the sample. Further detail about the test is available on our website[2].

There are separate one-hour tests for English and maths. They are based on the content and style for the new GCSEs in English language and maths. Each test is divided into several booklets. Each student takes just one booklet. We do not give each individual student a mark or a grade. We only want to measure overall performance across the whole sample of students that take the test.
  The test will provide additional information alongside existing evidence to indicate changes in performance nationally and so help us to maintain standards in the reformed GCSEs.

The National Reference Test in 2017

In 2017, we have used the results from the GCSEs in English language and maths to establish the scores in the NRT that equate to the GCSE grade boundaries. The NRT played no part in GCSE awarding in 2017. For example, we took the percentage of students in year 11 who achieved at least a grade 4 in the GCSE and determined the score in the NRT that the same percentage of students achieved. We did this for the grade boundaries at grade 4, 5 and 7 for both maths and English language.

---

[1] https://www.gov.uk/government/publications/an-overview-of-the-national-reference-test
[2] https://www.gov.uk/government/collections/national-reference-test-information
Ofoihihj0[q39jer[09 jwqo[eij

The purpose of the test in 2017 has therefore been to establish the baseline in standards in the NRT against which changes in performance in future years can be measured.

We needed a minimum of 300 schools to take part in the NRT. Ofqual's test supplier, NFER, prepared a stratified sample which included 360 schools in the 'main' sample. For each school in the main sample NFER also selected up to 3 replacement schools in case the selected school was not eligible or prepared to take part. Schools' mean GCSE grade in 2016 was used to construct the strata. All secondary schools in England were included in the population, except for very small schools and special schools. NFER approached 401 schools to take part, including 41 from the replacement sample.

We were very pleased that 341 schools agreed to take part, of which 314 were from the main sample of 360 and the remaining 27 being from the replacement school sample. Under legislation introduced in September 2016, most state schools were required to take part in the test, if asked to do so. However, the legislation does not apply to independent schools and a few academies. Of the 341 schools that took part, 325 were subject to the legislation and 16 agreed to do so voluntarily.

60 schools had been contacted (either in the main sample or the replacement) that did not take part. Only two of these were subject to the legislation. 19 schools were entering students only for an international GCSE and were therefore not eligible to take part. Four schools closed before the NRT was held. Other schools gave no reason why they would not take part or would not respond to the request.

Schools were asked to provide NFER with a list of their students in year 11 that were being entered for GCSEs English language and/or GCSE maths. NFER then selected a random sample of up to 30 students at each school to take the maths test and another 30 students to take the English test. (No student was asked to take both tests.) Of the 30 students for each subject, 24 took one of the live test versions. The other 6 students took a version of the test that included a combination of current live test items and some new items that we trialled and which we may choose to use in the live test in future years.

8,906 students took the maths test and 8,845 took the English test. 88% of students selected to take the test actually took part. The legislation allows a head teacher to withdraw an individual student from taking the test. This accounted for only about 1 in 20 of the absentees, indicating that heads were using this discretion only in exceptional circumstances. In most cases, the students who had been selected to take the test but who did not take part were absent from the school on the day of the test.

The tests took place in schools during a two-week period: 20th February to 3rd
Ofoihihj0[q39jer[09 jwqo[eij

March 2017. Schools selected the date and time when the test would take place within this period. Some chose to hold both subjects at the same time, others selected different times. Students had one hour to complete the test.

NFER assigned one or two test administrators to each school to make arrangements and to go to the school on the day of the test to administer it. The test administrators used a pre-agreed script to conduct the test to ensure consistency of approach in all schools.

Schools were able to request access arrangements for students taking the NRT, these being the same as the student would be offered when taking the GCSE. These included additional time, use of a word processor, a scribe or reader, enlarged test materials, test materials printed on coloured paper, a modified enlarged print version and Braille. Some schools requested that some of their students be accommodated in a separate room from the main group. One of NFER's test administrators was always present in each room. Only one test at one school did not take place as planned, caused by adverse weather which prevented the test administrator from travelling to the school.

NFER recruited and trained a total of 110 markers (including lead markers and team leaders), about half of whom had also marked the Preliminary Reference Test in 2016. Markers for the live test were trained on 4th March and completed the marking of the live booklets on schedule on 18th March. Extended response items in the English test were 100% double-marked with any out-of-tolerance items being blind-marked by a third marker.

NFER, together with CITO, the Dutch-based educational assessment and testing organisation, completed the analysis of the test data on schedule in May, using a psychometrics framework that had been agreed with Ofqual. This uses a combination of Classical Test Theory and Item Response Theory techniques[3]. Students only take a quarter of the complete test. The analysis enables individual students' responses to be combined to estimate a distribution of students' abilities. We have used the percentage of students who achieved at least a grade 4, a grade 5 and a grade 7 in their GCSE's this summer to establish the equivalent 'cut scores' on the distribution of abilities estimated in the NRT analysis. In future years, we will hold these 'cut scores' fixed and estimate the change in the percentage of students in the NRT that achieve them. This will provide the information on changes, if any, that may be expected in the distribution of grades in the GCSE awards.

We designed the NRT with the aim to achieve a level of test precision of ±1.5% (with a 95% confidence interval) at three grade boundaries: 7/6, 5/4 and 4/3. Table 1 shows the test precision achieved for the NRT in 2017.

---

[3] The software that was used is available under an open-source licence at : https://CRAN.R-project.org/package=dexter
Ofoihihj0[q39jer[09 jwqo[eij

Table 1: test precision in 2017

Subject / grade boundary	GCSE percentile (from summer awarding 2017)	NRT precision (percentage points)
English language		
- **Grades 4/3**	69.9%	±1.7
- **Grades 5/4**	53.3%	±1.8
- **Grades 7/6**	16.8%	±1.3
Maths		
- **Grades 4/3**	70.7%	±1.6
- **Grades 5/4**	49.7%	±1.7
- **Grades 7/6**	19.9%	±1.3

We concluded that the level of test precision achieved in 2017, if maintained at a similar level in subsequent years, is appropriate for the NRT results to be taken into account in the awarding of GCSEs.

The first time that we will have 2 sets of NRT data to compare will be in 2018. We would expect to see an improvement in early NRT results as students and teachers become more familiar with the new GCSEs in English language and maths. This, and the need to compare the performance of student cohorts in the NRT over several years, means information from the test will not be used by exam boards during awarding until at least 2019. From 2020, we will consider the evidence from the NRT each year when GCSEs are awarded.

Ofoihihj0[q39jer[09 jwqo[eij

We wish to make our publications widely accessible. Please contact us at publications@ofqual.gov.uk if you have any specific accessibility requirements.

OGL

Any enquiries regarding this publication should be sent to us at:

Office of Qualifications and Examinations Regulation

Spring Place
Coventry Business Park
Herald Avenue
Coventry CV5 6UB

Telephone   0300 303 3344
Textphone   0300 303 3345
Helpline   0300 303 3346