

Technical Information – New Group Maths Test (NGMT)

Trials

The questions for *New Group Maths Test* (NGMT) were developed by mathematics subject experts and trialling of these questions occurred in four rounds: summer 2018, winter 2018/19, autumn 2022 and summer 2023.

Across all trials, 16172 pupils from 200 schools were recruited. Although NGMT is aimed at ages 7-12 (Years Two-Seven in England), due to the adaptive nature of the assessment, material typically aimed at older pupils was trialled (against the appropriate ages) to minimise the prospects of the most capable pupils being adversely impacted by ceiling effects. Participation was distributed as follows:

Age (years) level	Number of students	Number of schools
<7	1550	53
7	1793	99
8	2374	121
9	2250	125
10	2351	115
11	2209	126
12	1939	56
13	1003	32
14	600	10
15+	102	7
Total	16171¹	200

In total, 1474 questions covering number, algebra, measurement, geometry, statistics, ratios, fractions, proportions and decimals were trialled. Due to the number of questions being developed, pupils could only see a subset of those available. Therefore, to ensure all questions were seen by enough pupils for reliable analyses to be conducted, five to seven forms were created for each Year Group such that between them they covered all of the appropriate material for the targeted age.

With the exception of the oldest pupils, the recruitment and form design outlined in the preceding text were successful in attaining the required samples for NGMT. Consequently, material targeted at 15 year olds could not be included in the assessment.

¹One pupil had no date of birth information provided.

Forms, both within and across Year Groups had shared questions enabling the relative difficulties of questions and abilities of pupils to be estimated and placed on a common scale which aided in determining if items were targeted at appropriate ages. The trials also provided information on:

- How well items discriminated between low and high scorers.
- If questions were equally difficult for boys and girls, after each gender's general ability had been controlled for.
- If questions were equally difficult, irrespective of the type of computing device used (desktop vs others), after controlling for the general ability of users.
- If questions that appeared across Year Groups were equally difficult for the older and younger pupils once each Year Group's general ability had been controlled for.

These analyses informed the construction of the question pool used in NGMT. The trials performed in 2018/19 also indicated, that as with NGRT and NGST, Year One pupils have difficulties using digital tests. Therefore, in accordance with the advice offered for the other members of the *New Group* suite of tests, we recommend that the youngest pupils assessed with NGMT should be Year Two.

Standardisation

Once the composition of the question pool had been defined, final difficulties for the questions could be determined. The difficulties were calculated via Item Response Theory, which also underpins the adaptive algorithm used to select questions during an administration of NGMT and allows an estimate of students' overall maths ability to be made. NGMT's standardised scores are based on the maths ability scores.

The standardisation for Form A was conducted in the United Kingdom from September-November 2023. A national database of schools was created and schools were grouped into categories by country (England, Wales, Scotland and Northern Ireland).

Typically, during a standardisation, schools within England would also be grouped on a measure of academic attainment. However, given the timing of the standardisation, the most recent data available for schools' academic performance came from 2019. Due to the amount of time that had elapsed since then and possible lingering effects due to Covid, it was decided to use the Income Deprivation Affecting Children Index (IDACI) as a proxy for attainment. The association between socio-economic status and attainment is well-established and IDACI had been calculated for 2023 and was publicly available by the time recruitment for the standardisation commenced ².

Schools were selected by stratified random sampling procedures within these groupings. As this was a national sample, many schools taking part in the standardisation had not used maths assessments from GL Assessment before.

²For subsequent NGMT forms, Key Stage 2 and GCSE data will be available and used for sampling.

In total, 13498 students from 101 schools took part in the standardisation and participation was broken down as follows:

Age (years) level	Number of students	Number of schools
6	1142	52
7	1687	58
8	1758	58
9	1776	58
10	2032	79
11	4520	86
12	583	34
Total	13498	101

Test reliability

The reliability of a test is a measure of the consistency of a student's test scores over repeated testing, assuming conditions remain the same – that is, there was no fatigue, learning effect or lack of motivation. Tests with poor reliability might result in very different scores for a student across two test administrations. The reliability of the test was estimated using the Cronbach's Alpha formula which produces values ranging from 0 to 1. Values above 0.70 are considered to be acceptable. The reliability of NGMT is 0.76.

For interpreting the score of an individual student, the Standard Error of Measurement (SEM) is a more useful statistic than a reliability coefficient. It indicates how large, on average, the fluctuations in standard scores may be and indicates the 68% chance or confidence band. However, tests typically show the 95% chance or confidence bands. The SEM for NGMT is 7.4 SAS points and for an average-performing student with a Standard Age Score (SAS) of 100, there is a 95% chance that the student's true SAS will be in the range +/- 15, i.e. from 85 to 115.

Gender differences

The tests have been age standardised to a national mean of 100 and standard deviation of 15. There were approximately similar numbers of males and females in the standardisation. Overall, males perform better than females by an average of 2.6 SAS points. The difference is statistically significant but small.

Maths Ability Scale

With an adaptive test, students within a class answer different questions depending on their abilities. Weaker students attempt easier questions whereas more able students attempt harder questions. Therefore, students may get the same number of questions correct but have different abilities. The ability scale score takes into account the difficulty of the questions a student has attempted and is the adaptive test equivalent to the raw score of a fixed test and as such fulfils the role raw scores do in a fixed test. The advantage with the NGMT scale score is you can use it to track the mathematical ability of an individual over time. SAS scores are adjusted for age and the average SAS score at any age group nationally is 100.

The chart and table below show the percentile distribution of the scale score by age. It shows that the average (50th percentile) scale score for students at age 6 years is 264 and increases to 339 at age 12. The top 10% of students (i.e. 90th percentile) at age 6 have a scale score of 335 or higher and at age 12 the top 10% of students have a scale score of 401 or higher.

