

# **2009 National curriculum tests review outcomes (provisional)**

---

November 2009

QCDA/09/4549

# Contents

Introduction .....	3
Reviews .....	3
Types of review .....	3
Review fees .....	5
Key figures for 2009.....	6
Technical information.....	8
Population of interest .....	8
Cohort numbers .....	9
Data sets.....	9
Reviews upheld.....	10
Rounding.....	10
Key to tables .....	10

## Introduction

This report provides provisional information on the outcomes of the reviews of the marking process for the 2009 national curriculum tests at key stage 2 in England. The figures in this report are produced from data provided by the test operations agency as part of the data feed to the Department for Children, Schools and Families (DCSF). The information in this report is provisional and is subject to the outcomes of a very small number of outstanding reviews of marking and process reviews.

## Reviews

A review is where a pupil's test script is checked to ensure that the original application of the *Mark scheme* was appropriate and that no clerical errors were made. A request for a review should be considered when, in the opinion of the staff in a school, the pupil(s) concerned have been awarded a national curriculum level above or below that to which their work is entitled in relation to the published *Mark scheme*.

## Types of review

There are three services available in 2009:

- Clerical check
- Individual review of marking
- Group review of marking (more than 30 pupils)

Individual and group reviews automatically include a clerical check as part of the service. Where an individual or group review request is not successful because the application of the *Mark scheme* by the original marker is deemed appropriate, but a clerical error is detected in the subsequent clerical check, the review is reported against a clerical check rather than the original review type.

The introduction of separate reading and writing markers for key stage 3 English in 2004 led to some consideration for accepting completely separate reviews for the reading and writing tests – although it is not a requirement for a component review system. In 2008, this was fully realised for both key stage 2 and key stage 3 English. For example, a school

may request a clerical check for English reading for one pupil where the marker had added the marks up incorrectly. The school may also include that pupil's English writing test script as part of a group review where they are unhappy about the quality of marking of the English writing marker. As, historically, review outcomes have been reported against the subject overall, in this case English, such a combination of review types doesn't naturally fall into one of the above three categories. Such reviews are coded as 'mixed' in the tables for 2008 and 2009.

Schools received their complete set of marked test scripts and pupil results by the published deadline in early July in order for them to make a decision about requesting reviews.

Until 2007, the primary mechanism for reporting results to schools and also collecting data nationally was a single paper marksheet, completed by markers. One copy of this marksheet was returned to schools along with the test scripts while another copy was sent to the data collection agency.

The marksheet contained attendance information completed by the school and marks for each test component transferred from the test scripts by the marker. Once this task was completed markers were required to calculate the total score for the subject, a simple sum of the component scores, and then convert the total mark to a level using a look-up table provided by QCDA. The data collection process simply captured the data as recorded by the marker.

In 2008, the new test operations agency moved away from this paper-based approach to an online mark capture system for markers. A benefit of this approach was that having captured the marks, the aggregations and level generation could be completed by computer reducing some of the administrative burden on markers, and also providing more accurate data by reducing the human error introduced during this routine task.

A consequence of the automated generation of the results for a school was that the results would be returned electronically from a central source rather than on paper by individual markers.

However, the particular implementation of the on-screen data capture process required markers to transfer question-level information rather than just the overall component scores. This was a significant increase in the burden on markers.

In 2009, the system of returning results to schools and collecting the national data was a mixture of the 2007 and 2008 approaches. The marksheets used in 2007 were reprinted and markers were required to transfer the component scores from test scripts to the marksheets. However, the aggregation and levelling parts of the process were completed by computer once the component scores had been collected. The paper marksheets, containing marks, were returned to schools along with their test scripts by the markers but schools were instructed to wait until they saw the official results on the *Pupil results* section of the *Test orders* website before submitting a review.

The group review is undertaken as a two part process. Initially a marking panel will consider the marking for a sample of pupils for whom the group review was submitted. If the marking for the sample reflects an appropriate application of the *Mark scheme*, no re-marking will be undertaken and the marks for the whole group will remain unchanged: 184 schools with 9989 pupils for English and one school with 93 pupils for science submitted review applications that fell within this category in 2009. The reviews data recorded this year have enabled this category to be reported separately.

## **Review fees**

Schools were informed that they would be charged for any reviews that are accepted that do not result in the change to the level (including reading and writing levels for English). The 2009 fees ranged from £5 for a clerical review to £6.50 for an individual review. For a group review, schools were charged a fee of £180 and for each pupil who moved up or down one or more levels, £6.50 was to be deducted from the fee (to a maximum of £180).

Once an individual or group review has been completed and the outcome reported to the school, if they are not satisfied that the correct procedures have been followed by the test operations agency, they may request a process review. The outcome of a process review is final and there is no right of appeal.

Details on the 2009 reviews process can be found on the QCDA *Tests and exams* website at <http://testsandexams.qcda.gov.uk/18959.aspx>.

## Key figures for 2009

- 1,676,315 pupil test scripts were marked in 2009.
- Review applications were made for a total of 50,257 pupil test scripts, which represents 3.0% of the total number of scripts marked.
- A total of 6,532 pupil test scripts received an overall subject level change (to a higher or lower level) as a result of the review request, representing a total of 0.39% of the total number of test scripts marked.

Key stage 2	Schools	Review requested†	Outcome	
			Lower level‡	Higher level‡
<b>English</b>				
Clerical review	1714	3281	147	946
		0.6%	4.5%	28.8%
Individual review	5105	27322	17	2562
		4.9%	0.1%	9.4%
Group review	73	4716	113	277
		0.9%	2.4%	5.9%
Group review§	184	9989	-	-
		1.8%	-	-
Mixed review	347	531	2	98
		0.1%	0.4%	18.5%
<b>Mathematics</b>				
Clerical review	238	277	5	236
		0.0%	1.8%	85.2%
Individual review	956	1201	0	626
		0.2%	0.0%	52.1%
Group review	1	96	2	1
		0.0%	2.1%	1.0%
Group review§	0	0	-	-
		-	-	-
Mixed review	0	0	0	0
		-	-	-

Science				
Clerical review	233	287	4	253
		0.1%	1.4%	88.2%
Individual review	1414	2464	0	1243
		0.4%	0.0%	50.4%
Group review	0	0	0	0
		-	-	-
Group review§	1	93	-	-
		0.0%	-	-
Mixed review	0	0	0	0
		-	-	-

† The percentage figures given in the 'Reviews requested' column use the cohort as the denominator in the calculation.

‡ The percentage figures given in the 'Outcome' columns use the count in the 'Reviews requested' column as the denominator in the calculation.

§ The group reviews, for which the application of the *Mark scheme* by the original marker was deemed to be correct.

In previous years, the data published on reviews of marking was based on the contractor's management information and not on the data feed provided to the DCSF. QCA has republished the data from 2006 using the information provided in the data feed alongside the 2009 data so that consistent business rules can be applied to all years. In addition, up to 2007, a standard cohort size of 650,000 was used to calculate percentages of the cohort that had applied for a review of marking and had a change of level as a result of a review of marking. QCDA has decided that actual cohort figures for each year should be used and has recalculated each percentage as appropriate.

From 2008, the marking process was different in some respects from previous years and this will make direct comparisons between years problematic. The borderlining process whereby pupils up to three marks below a level threshold had their test script reviewed by their original marker was removed in 2008. The reviews process in 2008 was also different from previous years in that all requests for review were accepted, regardless of

whether the correct procedures had been followed. In previous years, these requests would have been rejected and are not included in the figures.

QCDA does not believe that the outcomes of reviews of marking can be used to draw conclusions about the quality of marking in any year.

## **Technical information**

### **Population of interest**

The population of interest, or cohort, for each key stage and subject covers all schools in England with pupils participating in the end of key stage tests and who achieve a valid test outcome i.e. a national curriculum level 2, 3, 4, 5 or an award of 'N' (where too few marks are gained for the award of a level). In addition, there are a small number of Service Children's Education schools that are located overseas but who have pupils eligible for the end of key stage assessment who are included.

Pupils that do not sit the tests because of absence; because they are working below the level of the test; or because they are working at the level of the test but are unable to access them, are not included.

Pupils who sat one or more components of the test, but who do not have a full set of marks either through partial absence, test scripts being lost or because results have been annulled due to maladministration or malpractice, are not included.

In the tables that have school counts, schools submitting more than one type of review for a key stage or subject are counted separately for each type of review.



## Cohort numbers

The calculations of the types of review as a percentage of the cohort given in the table are based upon the following denominators, which are a count of pupils with national curriculum level 2, 3, 4, 5 or an award of 'N'.

Key stage 2	English	Mathematics	Science
2006	566,630	570,867	577,694
2007	559,523	563,080	571,024
2008	569,066	573,508	580,645
2009	554,219	557,841	564,255

Numbers for each subject for each year may vary for of a number of reasons, including:

- take up of the tests by independent schools
- absenteeism rates
- rates at which pupils make progress and complete the relevant programmes of study.

## Data sets

The data sets used are the ones provided by the data collection agency to DCSF as follows:

Year	Data feed	Date the data was provided
2006	5b	31 January 2007
2007	5	30 January 2008
2008	4k	6 October 2008
2009	6	16 October 2009

## **Reviews upheld**

Only reviews where the overall level for the subject changed as a consequence of the review are counted in the tables.

Any amendments to results outside the reviews process are not counted. This will include, for example, changes recorded by the DCSF contractor as part of the 2009 primary school Achievement and Attainment Tables – checking process with schools.

## **Rounding**

Any percentages given in this statistical release are given to one decimal place. The rounding convention is as follows: any fractions of 0.05 and above will be rounded up, anything less than 0.05 will be rounded down, for example, 4.483 will be rounded to 4.5; 4.445 will be rounded to 4.4.

## **Key to tables**

The following symbol notation is used in the tables published in this release:

- not applicable

~ not available.

## Trends over time

Key stage 2	2006				2007				2008				2009			
	Schools	Reviews requested†	Outcome: lower level‡	Outcome: higher level‡	Schools	Reviews requested†	Outcome: lower level‡	Outcome: higher level‡	Schools	Reviews requested†	Outcome: lower level‡	Outcome: higher level‡	Schools	Reviews requested†	Outcome: lower level‡	Outcome: higher level‡
English																
Clerical review	1036	1840	44	379	923	1460	47	357	431	914	29	209	1714	3281	147	946
		0.3%	2.4%	20.6%		0.3%	3.2%	24.5%		0.2%	3.2%	22.9%		0.6%	4.5%	28.8%
Individual review	1906	7262	5	507	784	1456	7	190	4630	25150	109	4103	5105	27322	17	2562
		1.3%	0.1%	7.0%		0.3%	0.5%	13.0%		4.4%	0.4%	16.3%		4.9%	0.1%	9.4%
Group review	58	2241	51	149	12	427	60	36	261	14781	243	736	73	4716	113	277
		0.4%	2.3%	6.6%		0.1%	14.1%	8.4%		2.6%	1.6%	5.0%		0.9%	2.4%	5.9%
Group review§	~	~	~	~	~	~	~	~	~	~	~	~	184	9989	-	-
		-	-	-		-	-	-		-	-	-		1.8%	-	-
Mixed review	~	~	~	~	~	~	~	~	94	296	7	63	347	531	2	98
		-	-	-		-	-	-		0.1%	2.4%	21.3%		0.1%	0.4%	18.5%
Mathematics																
Clerical review	363	475	13	77	336	430	12	80	262	359	22	235	238	277	5	236
		0.1%	2.7%	16.2%		0.1%	2.8%	18.6%		0.1%	6.1%	65.5%		0.0%	1.8%	85.2%
Individual review	581	753	0	325	332	373	0	292	1235	1720	0	941	956	1201	0	626
		0.1%	0.0%	43.2%		0.1%	0.0%	78.3%		0.3%	0.0%	54.7%		0.2%	0.0%	52.1%
Group review	1	65	1	2	0	0	0	0	4	126	1	0	1	96	2	1
		0.0%	1.5%	3.1%		-	-	-		0.0%	0.8%	0.0%		0.0%	2.1%	1.0%
Group review§	~	~	~	~	~	~	~	~	~	~	~	~	0	0	-	-
		-	-	-		-	-	-		-	-	-		-	-	-

Not protected

11

2009 National curriculum tests review outcomes (provisional)

Mixed review	~	~	~	~	~	~	~	~	2	2	0	1	0	0	0	0
	-	-	-	-	-	-	-	-		0.0%	0.0%	50.0%		-	-	-
Science																
Clerical review	336	422	12	64	318	390	9	82	260	338	11	240	233	287	4	253
		0.1%	2.8%	15.2%		0.1%	2.3%	21.0%		0.1%	3.3%	71.0%		0.1%	1.4%	88.2%
Individual review	741	1155	0	527	560	777	2	558	1645	2979	0	1765	1414	2464	0	1243
		0.2%	0.0%	45.6%		0.1%	0.3%	71.8%		0.5%	0.0%	59.2%		0.4%	0.0%	50.4%
Group review	0	0	0	0	0	0	0	0	4	242	0	0	0	0	0	0
	-	-	-	-	-	-	-	-		0.0%	0.0%	0.0%		-	-	-
Group review§	~	~	~	~	~	~	~	~	~	~	~	~	1	93	-	-
	-	-	-	-	-	-	-	-		-	-	-		0.0%	-	-
Mixed review	~	~	~	~	~	~	~	~	5	6	0	2	0	0	0	0
	-	-	-	-	-	-	-	-		0.0%	0.0%	33.3%		-	-	-

†The percentage figures given in the 'Reviews requested' column use the cohort as the denominator in the calculation.

‡The percentage figures given in the 'Outcome' columns use the count in the 'Reviews' requested' column as the denominator in the calculation.

§ The group reviews, for which the application of the *Mark scheme* by the original marker was deemed to be correct.