

Assessment of Achievement Programme

Sixth Survey of Mathematics, 2000

Findings and Issues

Scottish Executive Education Department

What is the Assessment of Achievement Programme?

The Assessment of Achievement Programme (AAP) was established by the Scottish Office Education and Industry Department in 1981 to monitor the performance of pupils in Scottish schools in particular areas of the curriculum. Since 1983, there have been regular surveys in three core curricular areas:

- English language
- mathematics
- science

The main objectives of the AAP are to:

- assess what pupils in P4, P7 and S2 know and can do
- provide information on performance in relation to the levels defined in the 5-14 National Guidelines
- provide evidence about changes in performance over time
- provide comparisons of performance and progression of pupils between P4 and P7 and P7 and S2
- provide comparison of performance between girls and boys

The surveys are intended to inform SEED, education authorities, teachers and other interested parties about the achievement of pupils.

The Sixth Survey of Mathematics

The sixth AAP survey of pupils' attainment in mathematics was carried out in 2000.

Samples of pupils at P4, P7 and S2 stages were selected to be representative of pupils in all mainstream schools. Around 9800 pupils completed the assessments.

Assessment was based on the curriculum defined in the National Guidelines: Mathematics 5-14. The reporting categories used in the survey were based on the following aspects of mathematics:

- Information handling
- Number, money and measurement
- Shape, position and movement
- Problem-solving

The assessment tasks were matched to Levels B – F as detailed in the National Guidelines.

The survey involved written assessments undertaken by all the pupils in the survey. Each pupil completed two test booklets. Each pair of booklets was carefully compiled to include questions covering all the strands being assessed in the survey at two levels for Primary 4 and three levels for Primary 7 and Secondary 2. There were 30 booklets at each stage. On average each booklet was completed by 200 pupils. This approach ensured that there was a very wide coverage of the mathematics syllabus – much wider than would be achieved in a single test taken by all pupils.

The survey also involved mental calculation tests taken by about a third of pupils in the survey.

Schools in the sample were invited to complete a school questionnaire about provision for mathematics at the relevant stages.

1. Performance in mathematics

For the 2000 mathematics survey, the results are presented in two formats: category (strand) scores and pupil scores.

Category (strand) scores

The **category (strand) scores** provide information about different attainment targets at different levels and highlight strengths and weaknesses in the various aspects of mathematics. The measure used was the mean percentage correct; that is the mean of all the pupil scores for all the items for a reporting category (strand) at a specific level e.g. *Add and Subtract* at Level D. This measure had been used in previous surveys.

To ensure a consistent basis for comments on the category (strand) scores, the following criteria were applied:

<i>Score</i>	<i>Comment</i>
80%	Very good - major strengths
66.7% - 79.9%	Good - strengths outweigh weaknesses
50% - 66.6%	Fair - some important weaknesses
< 50%	Unsatisfactory - major weaknesses

Primary 4

At Level B, the target level, the overall performance is good. It is very good in *Interpret* and in *Patterns and Sequences*, good in four of the other five categories and fair in *Multiply and Divide*. In no category is performance unsatisfactory.

Primary 7

At Level D, the target level, overall performance is fair. It is good in one category (*Add and Subtract*) and is fair in the other nine categories. In no category is performance unsatisfactory.

Secondary 2

At Level E, the target level, overall performance is fair. It is good in 1 category (*Add and Subtract*), fair in 7 categories and unsatisfactory in 1 category (*Fractions, Percentages and Ratio*).

Table 1: Mean percentage correct in each category (strand) at the target level for each stage

ASPECTS OF MATHEMATICS	REPORTING CATEGORY (Strand)	P4	P7	S2
		B	D	E
		Mean Score (%)	Mean Score (%)	Mean Score (%)
Information handling	<i>Interpret</i>	81.1	66.4	56.8
Number, money and measurement	<i>Range and Type of Number</i>	73.2	62.2	57.8
	<i>Patterns & Sequences</i>	84.5	55.2	54.8
	<i>Add & Subtract</i>	74.1	68.7	69.5
	<i>Multiply & Divide</i>	64.7	63.5	61.1
	<i>Fractions Percentages and Ratio</i>	-*	52.6	46.4
	<i>Time</i>	73.8	57.7	-*
Shape, position and movement	<i>Measure</i>	-*	58.1	57.7
	<i>Range of Shapes and Symmetry</i>	-*	63.6	53.0
	<i>Position & Movement and Angle</i>	73.4	52.8	60.0
Mathematics	Overall score	73.7	60.0	57.5

* - strand not assessed at this level

Pupil scores

Pupil scores are the percentage scores achieved by each pupil. Each pupil had a separate score for each level at which they were assessed. At P4, for example, pupils had separate scores for Level B and Level C. At P7 and S2 pupils had three scores relating to the three levels in their survey booklets.

To provide a wide picture of performance the pupil scores were recorded using three cut-off scores - 50%, 60% and 70%. For example, at P4, the percentages of pupils correct in 50%, 60% and 70% of the Level B items (and the percentages of pupils correct in 50%, 60% and 70% of the Level C items) are given.

For the purposes of reporting, the 60% cut-off score is used to judge whether pupils can be said to have attained a 5-14 level of attainment. Given the nature of the assessment, the proportion of pupils scoring more than 60% at the target level in the survey is considered to provide the best estimate of the percentage of pupils attaining that level.

Following the practice used in reports by HM Inspectors of Education, terms such as “most” and “few” are defined throughout in terms of the following percentages of the year group:

Almost all	over 90%
Most	75 - 90%
Majority	50 - 74%
Less (or fewer) than half	15 - 49%
Few	up to 15%

In P4, most pupils (77%) can be said to have attained Level B, the target level for their stage.

In P7 the majority of pupils (52%) can be said to have attained Level D, the target level for their stage.

In S2 almost half the pupils (49%) can be said to have attained Level E, the target level for their stage.

Table 2: The percentage of pupils achieving the cut-off score or above

Cut-off score	P4 Level B	P7 Level D	S2 Level E
50%	86.9	67.5	62.8
60%	77.3	52.0	49.4
70%	62.2	37.4	35.0

Performance over time

In order to find out what is happening to performance over time the results of the 2000 survey were compared with the results of previous survey in 1997. To make the comparison as valid as possible it was based only on tasks used in the 2000 survey which were also included in the 1997 survey. This comparison is not confined to the target level; it draws on all the common items available irrespective of level.

When this comparison is made the survey reveals that:

- performance at all three stages has improved significantly in most categories (strands)
- performance in P4 has improved significantly in all but one category (strand)
- the performance at P7 has improved significantly in all but two categories (strands)
- performance at S2 has improved significantly in all categories
- in no category (strand) has performance deteriorated.

**Table 3: Significance of performance in 2000 compared with performance in 1997
All common items**

REPORTING CATEGORY	P4 Levels B & C combined	P7 Levels C, D and E combined	S2 Levels D, E and F combined
	1997-2000	1997-2000	1997-2000
<i>Number Concepts</i>	Better	Better	Better
<i>Add and Subtract</i>	Better	Better	Better
<i>Multiply and Divide</i>	Better	No change	Better
<i>Fract. % & Ratio</i>	- *	Better	Better
<i>Measure inc. Time</i>	No change	-*	-*
<i>Time</i>	-*	Better	-*
<i>Measure</i>	-*	Better	Better
<i>Shape, Posn. & Movt.</i>	Better	No change	Better
Overall (all categories)	Better	Better	Better

* - strand not assessed at this level.

It is also possible to compare performance over a longer time, using the mean score on all the common items used at a stage.

Table 4: Mean percentage scores in all common items in surveys from 1988 to 2000

Year of Survey	1988	1991	1994	1997	2000
P4 (58 items)	58.7	58.8	56.0	59.0	63.9
P7 (63 items)	64.5	64.3	62.0	62.3	67.6
S2 (96 items)	58.3	57.1	55.8	56.4	62.2

From 1988 to 1994 changes were small with a dip in 1994 the first survey after the introduction of the National Guidelines which had set out a broad curriculum. There has been a significant improvement at all three stages from 1997 to 2000.

Performance between stages

In the survey some tasks were given to pupils in P4 and in P7 and some to pupils in P7 and S2. As a result it is possible to compare the performance of older pupils with the performance of younger pupils on the same tasks. For the P4/P7 comparison this was done on the Level C tasks; for the P7/S2 comparison this was done on the Level D and E tasks. The design of the survey allows the comparison to be done both on category (strand) scores and pupil scores. These comparisons are not measures of progression as different pupils were involved, but they do give an indication of progression.

Comparison using category (strand) scores (Table 5):

- There is noticeable progression in all categories from P4 to P7 with particular progress being made in *Fractions, Percentages and Ratio* and in *Position and Movement*.
- Progression in individual categories varies but the overall score shows a progression from 46.0 percentage points to 81.2 percentage points.
- From P7 to S2 there is a difference of about 10 percentage points in nearly all categories at level D.
- At Level E the progression is more variable.

Comparison using pupil scores (common items):

- From P4 to P7 there is a clear increase in the percentage of pupils attaining 60% of items at Level C. (Table 6)
- Only 5% of P7 pupils failed to score 50% on Level C items (the level below the target level). (Table 6)
- From P7 to S2 there is a clear increase in the percentage of pupils attaining 60% of items at both Level D and Level E. (Table 7)
- The proportion of pupils in S2 not attaining 50% at Level D, the level below the target level, is 18%. (Table 7) This is considerably larger than the equivalent proportion at P7 (5%) and gives cause for concern about lower attainers' learning in mathematics as they move into secondary school.

Table 5: Comparison of performance between stages: category (strand) scores

REPORTING CATEGORY	5 -14 LEVEL					
	C		D		E	
	Mean Score (%)		Mean Score (%)		Mean Score (%)	
	P4	P7	P7	S2	P7	S2
<i>Interpret</i>	63.4	88.5	66.4	76.4	42.1	56.8
<i>Range & Type of Number</i>	43.1	79.2	62.2	72.5	41.2	57.8
<i>Patterns & Sequences</i>	65.4	89.6	55.2	65.8	37.0	54.8
<i>Add & Subtract</i>	49.4	87.3	68.7	77.9	62.1	69.5
<i>Multiply & Divide</i>	41.3	84.2	63.5	71.2	53.8	61.1
<i>Fractions %s and Ratio</i>	23.5	76.5	52.6	62.9	35.1	46.4
<i>Time</i>	47.5	82.2	-	-	-	-
<i>Measure</i>	43.5	77.9	58.0	68.2	39.4	57.7
<i>Range of Shapes</i>	62.7	82.5	63.6	73.5	36.4	53.0
<i>Position and Movement</i>	23.4	67.8	52.8	70.2	29.9	60.0
Overall (all categories)	46.0	81.2	60.3	70.9	41.9	57.5

**Table 6: Comparison of performance between P4 and P7 in common assessment items –Level C
Percentage of pupils achieving cut-off scores**

Cut-off score	5-14 LEVEL	
	Level C	
	P4	P7
50%	44.1	94.7
60%	27.9	89.1
70%	16.2	79.4

**Table 7: Comparison of performance between P7 and S2 in common assessment items–Levels D and E
Percentage of pupils achieving cut-off scores**

Cut-off score	5-14 LEVEL			
	Level D		Level E	
	P7	S2	P7	S2
50%	67.5	82.1	36.8	62.8
60%	52.5	70.7	23.3	49.4
70%	37.4	57.2	13.1	35.0

2. Mental Calculation

Pupils' performance in mental calculation was assessed in the 2000 survey, as it had been in the 1997 survey. A subset of pupils taking the written assessments also sat a mental calculation test. There were five tests at each stage. At P4 the tasks were all at the target level, at P7 and S2 half the tasks were at the target level and half at the level below. Questions mainly involved the four basic operations – add, subtract, multiply and divide. The following are examples of typical tasks:

Primary 4 *Add 26 and 5. (94.4% answered correctly)*

Primary 7 *Subtract 15 from 49. (86.5% answered correctly)*

Secondary 2 *Multiply 9 times 30. (87.6% answered correctly)*

Each task was read out to the pupils, then repeated. Pupils had a short time to write down the answer before the next task was read out.

Scores

Table 8 shows the mean score at each stage based on items at the target level.

Table 8: Mean score at each stage based on items at the target level

P4 Level B	P7 Level D	S2 Level E
79.8%	74.3%	65.9%

Table 9 shows the percentage of pupils who can be said to have attained the target level when a cut-off score of 60% is applied to the results.

Table 9: Percentage of pupils at each stage achieving the cut-off score or above

P4 Level B	P7 Level D	S2 Level E
86.7%	78.9%	65.5%

Comparison over time

Many of the items used in 2000 had been used in the 1997 survey. A comparison in performance can therefore be made using these common items. The results of this comparison are shown in Table 10

Table 10: Mean scores in mental calculation on common items used in 1997 and 2000

	Mean % Score	
	1997	2000
P4	75.3	83.0
P7	81.2	80.9
S2	68.0	77.6

At P4 and at S2 percentage scores have increased impressively. At P7 scores in 1997 were already high and there is hardly any change in percentage scores in 2000.

3. Problem-solving

Problem-solving is an important aspect of mathematics. Short response problem-solving tasks were included in the survey at all three stages. A task in the survey was considered to be a “problem” if it satisfied one or more of the following criteria:

- more than one operation was necessary to solve the problem
- the operations(s) to be used was (were) not immediately obvious
- some interpretation was necessary.

The problems were assessed on the basis of ‘correct answer’ or ‘incorrect answer’. There was no attempt to look separately at the three components of problem-solving - ‘starting the task’, ‘doing the task’ and ‘reporting the task’. Pupil performance in problem-solving in the survey, therefore, is not an assessment of pupil performance as defined in the attainment outcome *Problem-solving and enquiry*. There are no 5-14 levels assigned to Problem-solving and enquiry. For the purposes of the AAP survey, levels were assigned to the short problems normally according to the level of the mathematics contained in the problem. For a small number of items the level assigned was the level above the level of the mathematics contained in the problem.

Table 11 Mean scores in problem-solving at the target level

Stage/level	Mean score (%)
P4 Level B	39.8
P7 Level D	48.8
S2 Level E	49.1

Problem-solving results should be treated with some caution. At each of the three stages P4, P7 and S2 problem-solving results are significantly lower than results in other categories. Problem-solving items usually contain more words and may require comprehension and interpretative skills as well as mathematical ones. They also often contain more than one mathematical skill.

The results indicate that problem-solving is an area of mathematics which continues to need attention. Particular attention may be needed to the language of mathematics.

4. Gender comparison

Using the pupil scores with a 60% cut-off score:

- at P4 and P7 there was no significant difference in the overall performance of boys and girls in mathematics
- at S2 girls performed overall better than boys (but boys and girls are equally represented amongst the lowest attainers)..

In the different aspects of mathematics:

- girls in S2 performed significantly better than boys in *Add and Subtract* and *Multiply and Divide*.
- at P4 and S2 there was no significant difference in the performance of boys and girls in mental calculation or problem-solving
- at P7 boys performed better than girls in mental calculation.

5. Information from schools

Schools participating in the survey were asked to complete a questionnaire about the provision of mathematics.

- Almost all schools taught mathematics for at least the recommended minimum time per week. Primary schools taught mathematics for longer each week than in 1997.
- Secondary schools were moving away from mixed ability classes to setting or broad banding at S2.
- Primary classes were mainly taught in groups but there had been a movement towards whole class teaching since 1997. Similarly, secondary teachers were also spending more time teaching the whole class.
- Approximately half of the primary schools and nearly three quarters of secondary schools had a policy for the use of calculators.
- Mental calculation was given daily in most primary schools and more than once a week in the majority of S2 classes.
- A wide variety of strategies for mental calculation were used with primary pupils.
- Computers for teaching and learning mathematics were more readily available in primary schools than secondary schools.

6. Conclusions and issues

The performance of pupils at P4, P7 and S2 has improved significantly since the previous survey in 1997. The performance of pupils at all three stages has also improved significantly when compared to the average scores over the previous four surveys going back to 1988. Results therefore are positive and show a welcome rise in standards of attainment in mathematics in Scotland.

However, performance in the past has not been good and performance in 2000 is still not very good, particularly at P7 and S2. The use of pupil scores for the first time underlines this. Using a test cut off score of 60% as an indicator of attainment of a level, 77% of P4 pupils can be said to have attained their target level, but only 52% of P7 pupils and 49% of S2 pupils do so. The improvement indicated by the survey will need to be sustained and bettered to raise attainment to a more satisfactory level.

The findings of this survey raise issues that should be helpful in the ongoing debate by teachers as to how the levels of attainment in mathematics in our schools can be improved.

- Have the changes to classroom practice, such as increased attention to mental approaches, influenced the rise in attainment in both mental calculation and possibly other areas of mathematics? Are there other aspects of classroom practice that might bring about increases in attainment?
- What more can be done in the learning and teaching of fractions and percentages to help pupils with these topics which, at all stages, they find most difficult?
- It is often difficult to tell if a child's difficulty is a mathematical one or a language one. Is it with the mathematical concept, the understanding of the problem in context, the language of mathematics itself or just the ability to read the question? All are issues to be addressed in the learning and teaching of mathematics.
- Aspects in all strands, even the ones with the best scores, cause pupils problems. Are school mathematics programmes addressing these topics in sufficient breadth and depth or returning to them often enough?