Technical Information

PT MATHS

The questions for *Progress Test in Maths* (*PTM*) were developed by the Mathematics Assessment Resource Service (MARS) team at the University of Nottingham. For each test level of the paper version, around 50% more questions were developed than required for the final version of the test booklet.

All test items were then trialled and standardised and final test items selected. The following explains this process first for levels 5-11T and then for 12-14, including information on reliability and comparison tables for the previous edition of the test.

*PTM*5-11T

Test questions were trialled in spring 2013 using three test paper booklets for each year group with common questions between booklets. Each question was taken by around 250 students. The numbers of students taking part in the trials were as follows.

Test level	Number of students	
	Paper	Digital
PTM5	750	
PTM6	776	
PTM7	764	270
PTM8	836	453
PTM9	629	599
PTM10	689	489
PTM11	857	395
PTM11T	606	0
Total	**5907**	**2206**

The data from the paper trials were analysed to provide information on the difficulty level of each question, its ability to discriminate between high and low scorers, and the extent to which it proved equally difficult for both genders, once each gender's general level of performance was taken into account. This information was then used to select questions for the final standardisation version of the paper and digital versions. A few of the questions were modified to enable these to work in the digital mode and the digital versions were trialled in spring 2014.

Standardisation

Two separate standardisations were conducted: one for the United Kingdom and one for Northern Ireland. Sufficient numbers of students were required for each

of these standardisations. The standardisation of the paper version of *PTM* took place between February and May 2014 for *PTM*5-11 and between September and October 2014 for *PTM*11T. A national database of schools was created and schools were grouped into categories by country (Wales, Scotland and Northern Ireland). In England, schools were further grouped into independent or grammar plus five categories of school intake based on overall school performance at the end of primary schooling using Key Stage 2 outcomes, or for secondary schools using the GCSE outcomes. For the Northern Ireland standardisations, schools were categorised into independent or grammar plus five categories of school intake based on uptake of free school meals within the schools.

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 never used maths assessments from GL Assessment before. For the standardisation, schools were asked to do one pre-selected *PTM* test level and were given an option to do other levels. Primary schools were asked to test all students in the year group but secondary schools had the option to either test two randomly selected teaching groups or the whole year group.

The numbers of students taking part in the two standardisations were as follows.

Test level	Number of students	
	United Kingdom*	Northern Ireland
PTM5	2786	810
PTM6	3335	1132
PTM7	4071	1408
PTM8	4609	1573
PTM9	3354	856
PTM10	4497	1290
PTM11	4823	1738
PTM11T	7287	1955
Total	**34762**	**10762**

* The UK numbers include the Northern Ireland numbers.

Schools were free to choose between the paper and digital version of the test and most primary schools chose the paper version. Overall, around 7137 students completed the digital version and 27,625 students did the paper version.
The students in the Northern Ireland standardisations were included in the UK standardisations but are over-represented in the UK sample compared with the proportions nationally. The results were weighted to reflect the national distributions of students in the categories mentioned above separately for the paper version and the digital version.

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 0.80 are considered to be very good. The reliability values for the various *PTM* batteries are given in the table below, separately for the UK and the Northern Ireland standardisations, and all show that the tests are very reliable. There were no significant differences between the reliabilities between the two regions, and further analysis showed that the reliabilities for the paper and digital versions were very similar.

Test level	Reliability	
	UK	NI
PTM5	0.87	0.85
PTM6	0.90	0.89
PTM7	0.91	0.91
PTM8	0.92	0.92
PTM9	0.94	0.94
PTM10	0.95	0.95
PTM11	0.95	0.96
PTM11T	0.95	0.94

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, most tests show the 90% chance or confidence bands. The SEM for *PTM*10 is 3.5, and for an average-performing student with a *PTM*10 Standard Age Score (SAS) of 100, there is a 90% chance that the student's true SAS will be in the range +/- 6.0, i.e. from 94 to 106.

Test level	SEM		90% SAS confidence band (+/-)	
	UK	NI	UK	NI
PTM5	5.4	5.8	9	10
PTM6	4.7	5.0	8	8
PTM7	4.5	4.6	7	8
PTM8	4.2	4.3	7	7
PTM9	3.8	3.6	6	6
PTM10	3.5	3.5	6	6
PTM11	3.3	3.1	5	5
PTM11T	3.4	3.7	6	6

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 standardisations. The table below shows the mean SAS score differences between males and females for UK and for Northern Ireland. The largest discrepancy is for *PTM*5 and *PTM*11T where there is an average difference of around 2 points for SAS scores. In most cases the differences are small and within +/- 1 SAS point.

Test level	Gender - mean SAS differences	
	UK standardisation	NI standardisation
PTM5	2.3	1.7
PTM6	-0.3	0.3
PTM7	-0.5	0.8
PTM8	-0.2	-0.2
PTM9	-1.3	-1.0
PTM10	-1.5	0.5
PTM11	-0.5	0.7
PTM11T	-2.4	1.7

Note - positive scores: females higher than males; negative scores: females lower than males.

Northern Ireland compared with UK

The large numbers of students taking part in the separate Northern Ireland standardisations enable us to compare students in Northern Ireland with students in the UK. The table below shows the average SAS differences based on the UK standardisations with a UK mean SAS of 100. The differences are small.

Test level	Mean SAS difference
PTM5	-1.0
PTM6	0.0
PTM7	0.1
PTM8	1.0
PTM9	-0.1
PTM10	1.7
PTM11	1.4
PTM11T	0.5

Note - positive scores: Northern Ireland higher than UK; negative scores: Northern Ireland lower than UK.

Relationship between *PTM* and *PIM* outcomes

PTM has been designed to replace the well-established series, *Progress in Maths* (*PIM*). A comparative study, in which about 350 students from each year group took both the old and the new versions of the tests was undertaken for each pair of tests from the two series. A sample of students aged 5 to 11 years old had to take both the new and older version of the test. Each pair of tests was then statistically equated, so that each standardised score on the old test could be mapped onto an equivalent score on the new test.

The strength of the relationship between two variables can be measured by a statistic called the correlation coefficient. A value of zero indicates no relationship between the two measures whereas a value of one indicates a perfect positive relationship. The table below shows the correlations for each level and that the correlations are all highly significant.

Test level	Correlation
PTM5	0.62
PTM6	0.78
PTM7	0.80
PTM8	0.79
PTM9	0.85
PTM10	0.86
PTM11	0.88

Schools that have data on students' standardised scores from *PIM* will clearly wish to convert these so that they can be compared directly to the students' results on the new tests. The conversion table on the next pages gives this information. Teachers should take the SAS from the old test series and convert this to the equivalent at the same level of *PTM*. So, for example, a standardised score of 95 on *Progress in Maths* 8 is equivalent to a standardised score of 96 on the new *Progress Test in Maths* 8. This is the starting point for making a judgement by comparing this with the SAS from the next level of *PTM*. From the example above, this year's *PTM9* score should be compared with last year's *PIM* score of 95 that has been converted to 96.

Progress in Maths SAS	Equivalent Progress Test in Maths SAS						
	PTM5	PTM6	PTM7	PTM8	PTM9	PTM10	PTM11
69	78	73	70	73	69	69	69
70	79	73	71	73	69	69	69
71	80	74	72	74	69	69	69
72	81	75	73	75	69	70	69
73	82	76	74	76	69	71	70
74	82	77	75	77	70	72	71
75	83	78	76	78	72	73	72
76	84	79	77	79	73	74	73
77	85	79	78	80	74	75	74
78	86	80	79	81	75	76	75
79	87	81	80	82	76	77	76
80	88	82	81	83	77	78	77
81	88	83	82	83	78	79	78
82	89	84	83	84	79	80	79
83	90	85	84	85	80	81	80
84	91	85	85	86	81	82	81
85	92	86	86	87	82	83	83
86	93	87	87	88	83	84	84
87	94	88	88	89	84	85	85
88	94	89	89	90	85	86	86
89	95	90	90	91	86	87	87
90	96	91	91	92	87	88	88
91	97	92	92	93	88	89	89
92	98	92	93	93	89	90	90
93	99	93	94	94	90	91	91
94	100	94	95	95	91	92	92
95	101	95	96	96	93	93	93
96	101	96	96	97	94	94	94
97	102	97	97	98	95	95	95
98	103	98	98	99	96	96	96
99	104	98	99	100	97	98	97
100	105	99	100	101	98	99	98

Progress in Maths SAS	Equivalent Progress Test in Maths SAS						
	PTM5	PTM6	PTM7	PTM8	PTM9	PTM10	PTM11
101	106	100	101	102	99	100	99
102	107	101	102	103	100	101	100
103	107	102	103	103	101	102	101
104	108	103	104	104	102	103	102
105	109	104	105	105	103	104	103
106	110	104	106	106	104	105	104
107	111	105	107	107	105	106	105
108	112	106	108	108	106	107	106
109	113	107	109	109	107	108	107
110	113	108	110	110	108	109	108
111	114	109	111	111	109	110	109
112	115	110	112	112	110	111	110
113	116	110	113	113	111	112	111
114	117	111	114	113	112	113	112
115	118	112	115	114	114	114	113
116	119	113	116	115	115	115	114
117	119	114	117	116	116	116	115
118	120	115	118	117	117	117	116
119	121	116	119	118	118	118	117
120	122	117	120	119	119	119	118
121	123	117	121	120	120	120	119
122	124	118	121	121	121	121	120
123	125	119	122	122	122	122	121
124	125	120	123	123	123	123	122
125	126	121	124	123	124	124	123
126	127	122	125	124	125	125	124
127	128	123	126	125	126	126	125
128	129	123	127	126	127	127	126
129	130	124	128	127	128	128	127
130	131	125	129	128	129	130	128
131	132	126	130	129	130	131	129
132	132	127	131	130	131	132	130

Progress in Maths SAS	Equivalent Progress Test in Maths SAS						
	PTM5	PTM6	PTM7	PTM8	PTM9	PTM10	PTM11
133	133	128	132	131	132	133	131
134	134	129	133	132	134	134	132
135	135	129	134	133	135	135	133
136	136	130	135	134	136	136	134
137	137	131	136	134	137	137	135
138	138	132	137	135	138	138	136
139	138	133	138	136	139	139	137
140	139	134	139	137	140	140	138
141	140	135	140	138	141	141	141

*PTM*12-14

Test questions were trialled between January and May 2014 using three test paper booklets for each year group with common questions between booklets. Each question was taken by around 300 students. The numbers of students taking part in the trials were as follows.

Test level	Number of students
PTM12	1186
PTM13	1246
PTM14	1002
Total	**5907**

The data from the trials were analysed to provide information on the difficulty level of each question, its ability to discriminate between high and low scorers, and the extent to which it proved equally difficult for both genders, once each gender's general level of performance was taken into account.

This information was then used to select questions for the final standardisation version of the paper and digital versions. A few of the questions were modified to enable these to work in the digital mode.

Standardisation

The standardisation of *PTM*12-14 took place between February and May 2015. A national database of schools was created and schools were grouped into categories by country (Wales, Scotland and Northern Ireland). In England, schools were further grouped into independent or grammar plus five categories of school intake based on overall school performance for secondary schools using the GCSE outcomes.

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 never used maths assessments from GL Assessment before. For the standardisation, schools were asked to do one pre-selected *PTM* test level and were given an option to do other levels. Secondary schools had the option to either test two randomly selected teaching groups or the whole year group.

The numbers of students taking part in the standardisations were as follows.

Test level	Number of students
PTM12	4386
PTM13	4122
PTM14	4191
Total	**12699**

Schools were free to choose between the paper and digital version of the test and most secondary schools chose the paper version. Overall, around 21% students completed the digital version and 79% students did the paper version. The results were weighted to reflect the national distributions of students in the categories mentioned above separately for the paper version and the digital version.

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 0.80 are considered to be very good. The reliability values for the various *PTM* batteries are given in the table below and all show that the tests are very reliable. There were no significant differences between the reliabilities between the paper and digital versions.

Test level	Reliability
PTM12	0.95
PTM13	0.94
PTM14	0.94

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, most tests show the 90% chance or confidence bands. The SEM for *PTM*12 is 3.4, and for an average-performing student with a *PTM*10 Standard Age Score (SAS) of 100, there is a 90% chance that the student's true SAS will be in the range +/- 6.0, i.e. from 94 to 106.

Test level	SEM	90% SAS confidence band (+/-)
PTM12	3.4	6
PTM13	3.7	6
PTM14	3.7	6

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 standardisations. The table below shows the mean SAS score differences between males and females. The difference is small for *PTM*12 but scores for males are on average 2 points higher than for females for *PTM*13 and *PTM*14.

Test level	Gender - mean SAS differences
PTM12	0.5
PTM13	-1.7
PTM14	-2.2

Note - positive scores: females higher than males; negative scores: females lower than males.

Relationship between *PTM* and *PIM* outcomes

PTM has been designed to replace the well-established series, *Progress in Maths* (*PIM*). A comparative study, in which about 250 students from each year group took both the old and the new versions of the tests was undertaken for each pair of tests from the two series. Each student had to take both the new and older version of the test. Each pair of tests was then statistically equated, so that each standardised score on the old test could be mapped onto an equivalent score on the new test. The strength of the relationship between two variables can be measured by a statistic called the correlation coefficient. A value of zero indicates no relationship between the two measures whereas a value of one indicates a perfect positive relationship. The table below shows the correlations for each level and that the correlations are all highly significant.

Test level	Correlation
PTM12	0.87
PTM13	0.83
PTM14	0.87

Schools that have data on students' standardised scores from *PIM* will wish to convert these so that they can be compared directly to the students' results on the new tests. The conversion table on the next page gives this information.

Teachers should take the SAS from the old test series and convert this to the equivalent at the same level of *PTM*. So, for example, a standardised score of 95 on *Progress in Maths* 13 is equivalent to a standardised score of 93 on the new *Progress Test in Maths* 13.

Progress in Maths Score	Equivalent Progress Test in Maths SAS		
	PTM12	PTM13	PTM14
69	69	71	69
70	69	72	69
71	69	73	69
72	69	74	70
73	69	75	71
74	69	76	72
75	69	76	73
76	70	77	74
77	71	78	75
78	72	79	76
79	73	80	77
80	74	81	79
81	75	81	80
82	76	82	81
83	77	83	82
84	78	84	83
85	79	85	84
86	80	85	85
87	81	86	86
88	82	87	87
89	83	88	88
90	84	89	89
91	85	89	90
92	86	90	92
93	87	91	93
94	88	92	94
95	89	93	95
96	90	94	96
97	91	94	97
98	92	95	98
99	93	96	99
100	94	97	100
101	95	98	101
102	96	98	102
103	97	99	104
104	98	100	105
105	99	101	106

Progress in Maths Score	Equivalent Progress Test in Maths SAS		
	PTM12	PTM13	PTM14
106	100	102	107
107	101	103	108
108	102	103	109
109	103	104	110
110	104	105	111
111	105	106	112
112	106	107	113
113	107	108	114
114	108	109	116
115	109	110	117
116	110	111	118
117	111	112	119
118	112	113	120
119	113	114	121
120	114	115	122
121	115	116	123
122	116	117	124
123	117	118	125
124	118	119	126
125	119	120	128
126	120	121	129
127	121	122	130
128	122	123	131
129	123	124	132
130	124	125	133
131	125	126	134
132	126	127	135
133	127	128	136
134	128	129	137
135	129	130	138
136	130	131	139
137	131	132	141
138	132	133	140
139	133	134	141
140	134	135	141
141	135	136	141