READING PROGRESS TESTS, Stage 2
Australian Norms Supplement (Revised, November 2001)

Prepared by Marion M. de Lemos for ACER Press

The *Reading Progress Tests*[1] are a series of British tests covering the age range from 5 to 11 years. The series comprises the *Literacy Baseline* test, designed for use at the beginning of the first year of formal schooling in England and Wales (Year 1), and *Reading Progress Tests 1 to 6*, designed for use at the end of each of the six years of primary schooling (Years 1 to 6). The *Literacy Baseline* test assesses pre-reading and early reading skills, while *Reading Progress Tests 1 to 6* cover children's comprehension of written text at increasing levels of difficulty.

These tests have filled a need in the Australian market for measures of reading which can be used to assess children's reading skills from the first year of school to the end of the primary level. The *Literacy Baseline* test is designed to be administered either individually or in small groups, while the six *Progress Tests* are group administered and are designed to be used with whole class groups. The tests do not have set time limits, but each test would normally take about 45 to 50 minutes to administer.

These tests provide a valid measure of early literacy and reading comprehension at the primary level, and are suitable in content and format for use in Australia. However, the British normative data are not applicable in the Australian context, and this has led to difficulties in the interpretation of results. There are two main reasons for these difficulties.

The first difficulty relates to the mismatch in age between the six primary year levels in the British system and the corresponding year levels in the Australian system, which means that Australian children are on average eight to nine months older than children at the corresponding year level in the United Kingdom, except in the case of those states which do not have a pre-Year 1 year prior to entry to Year 1, currently Queensland and Western Australia.

The second difficulty, which is related to this age mismatch, is that the British normative data are presented in the form of age norms. This means that, when the tests are used at the end of the school year, up to a half or more of the children at the corresponding year level in Australia fall outside the age range of the age norms provided.

A further difficulty in the use of the British age norms, and particularly the reading ages corresponding to specific raw scores, is the fact that these age norms are based on the assumption of a linear relationship between age and test score. In the case of the Australian samples, there is relatively little increase in test score with age within a year level, except in the first or second year of schooling, and the use of age norms which assume such a relationship therefore give misleading results when applied in the Australian context. Because the standardised score corresponding to a specific raw score is lower for older children than for younger children with the same raw score, the use of the British age norms poses particular problems when comparing the relative performance of older and younger children within a year level. This problem applies particularly to the use of reading ages for the interpretation of raw scores, since outside a fairly narrow band of raw scores the reading age equivalent is based on extrapolation from the real data. The British Manual points to some of the problems associated with the construction of reading ages and the need to interpret reading ages with caution. This caution applies even more strongly in the Australian context, where the mismatch with the British age /grade relationship combined with the lack of any consistent increase in score with age could lead to misleading interpretation of results, particularly when scores are converted to reading ages that fall outside the chronological age range of the children who provided the data on which the tables are based.

Given these difficulties in using the British norms for the interpretation of Australian scores on the *Reading Progress Tests* it was decided to provide Australian users with normative data based on Australian samples. In the case of the Stage 1 series of tests (the *Literacy Baseline* test, *Reading Progress Test 1* and *Reading Progress Test 2*),

[1] Published by Hodder and Stoughton, 1996

Australian norms were constructed on the basis of data collected as a part of the ACER Project on *Curriculum and Organisation in the Early Years of School (see Australian Supplement for the Stage 1 tests).* In the case of the Stage 2 tests, Australian data on the *Reading Progress Tests* were collected as a part of the norming program for the revised *Progressive Achievement Tests (PAT) Reading Comprehension* and *Reading Vocabulary.*

Australian Norms for Reading Progress Tests 3 to 6

The norming of the revised *PAT Reading Comprehension* and *Reading Vocabulary* was undertaken in November 1999, and the study was designed in such a way that data were collected on several different tests at the same time. A summary of the norming design, and the tests administered at each level of schooling, is shown in Table 1.

Table 1 Summary of the Norming Design for the Restandardisation of the Revised PAT Reading and Associated Tests

Year Level	First Session All Groups	Second Session Group A	Second Session Group B	Second Session Group C
3	PAT Reading Comprehension and Reading Vocabulary	TORCH[2]	Reading Progress Test 3	Group Reading Test, Forms A and B
4	PAT Reading Comprehension and Reading Vocabulary	TORCH	Reading Progress Test 4	Group Reading Test, Forms A and B
5	PAT Reading Comprehension and Reading Vocabulary	TORCH	Reading Progress Test 5	Group Reading Test, Forms A, B, C and D
6	PAT Reading Comprehension and Reading Vocabulary	TORCH	Reading Progress Test 6	Group Reading Test, Forms A, B, C and D
7	PAT Reading Comprehension and Reading Vocabulary	TORCH	Group Reading Test Forms X and Y	Group Reading Test, Forms A, B, C and D
8	PAT Reading Comprehension and Reading Vocabulary	TORCH	Group Reading Test Forms X and Y	Group Reading Test, Forms C and D
9	PAT Reading Comprehension and Reading Vocabulary	TORCH	Group Reading Test Forms X and Y	Group Reading Test, Forms C and D

The Norming Sample

A nationally representative sample of 100 primary schools and 100 secondary schools was selected to participate in the norming of the *PAT Reading* series of tests. This sample was selected on a probability proportional to size basis, and included State, Catholic and Independent schools, with representation according to both State and type of school proportional to their representation in the total school population. The sample included all states and territories except the Northern Territory, which was not included in the sampling design. However, not all of the schools approached agreed to participate in the study, and of those which did agree to participate not all completed the testing. The final achieved sample therefore was somewhat less than the original target sample. The final sample comprised a total of 56 primary schools and 30 secondary schools. Of these, 21 primary schools (in Sample Group B) participated in the collection of data on the *Reading Progress Tests 3 to 6.*

[2] TORCH Tests of Reading Comprehension

Comparison of British and Australian Results on Reading Progress Tests 3 to 6

An initial analysis of the results on the *ReadingProgress Tests 3 to 6* was undertaken to examine the mean scores of Australian students on each test at the relevant year level, and at the same time to compare the results of the Australian norm group with that of the British norm group according to year of schooling and according to age. The results of these analyses are shown in Table 2.

In making comparisons with the British data, it should be noted that the British year levels (in England and Wales) are in general equivalent to the Australian year levels, although children in Britain start school at a younger age than children in Australia, leading to a difference of approximately eight months in average age between Australian and British children at equivalent year levels.

Table 2 Comparison of Australian and British Mean Scores at each Level

Data on British Norm Sample					Data on Australian Norm Sample				Comparisons with British Norms			
Test	Year Level (B & A)[3]	N	Mean Age	Mean Raw Score	N	Mean Age	Mean Raw Score	Mean PAT RC SS[4]	Equiv Age SS[5]	Equiv Grade SS[6]	Equiv Ability Score[7]	Equiv Reading Age[8]
RPT 3	3	2014	8:4	19.9	442	9:0	23.7	101.6	102	107	97	9:9
RPT 4	4	2540	9:4	23.2	410	10:0	23.9	100.6	95	101	100	9:4
RPT 5	5	2702	10:4	30.4	381	11:0	29.5	99.7	93	100	102	10:2
RPT 6	6	2424	11:3	34.6	271	12:1	35.4	100.9	94	100	104	11:2

For each of the tests listed in the left hand column (RPT 3 to 6), information is provided on the data from both the British and the Australian samples. The first column indicates the year level at which the test was administered, while the following three columns provide information on the data from the British sample (the size of the sample, the mean age of the sample, and the mean raw score on each test). The next three (shaded) columns provide comparable information on the Australian sample. The last four columns in the table provide an indication of how the Australian results would be interpreted in terms of the British norms. The first of these columns provides the age standardised score, based on the British norms, that would be achieved for a raw score at the age corresponding to the mean age of the group. The second column provides an estimated grade standardised score, based on the British norms for the age group obtaining a mean standardised score of 100 on the basis of the British norms. The third and fourth columns indicate the ability score and the reading age corresponding to the mean raw score of the Australian sample, using the conversion tables provided in the British Manual.

In order to check whether the subsamples of Australian students who took the *Reading Progress Tests* is comparable, in terms of reading ability, with the total Australian sample at each year level, the mean standardised score of these subsamples on the *PAT Reading Comprehension* test (PAT RC SS) is also indicated. This score is shown in the last of the shaded columns, that is, the column after the column showing the mean raw score of the Australian sample on each test. As can be seen from the figures in this column, the mean standardised score on the *PAT Reading Comprehension* test of the samples of students who also did the *Reading Progress Tests* is close to the average in each case, thus indicating that these samples are representative samples of Australian children in terms of reading ability, as assessed on the *PAT Reading Comprehension* test. The data from these samples can therefore be taken as valid indicators of the performance of Australian children on this series of British tests.

[3] Both Australian and British norms are based on testing undertaken at the end of the school year (November/December in Australia, June in the United Kingdom (England and Wales).

[4] Mean standardised score on the *PAT Reading Comprehension* test for the Australian Reading Progress Test samples

[5] Equivalent age standardised score for Australian sample based on British norms for age group corresponding to mean age of Australian sample.

[6] Equivalent grade standardised score for Australian sample based on British norms for British age group obtaining a mean score of 100 on the basis of the British norms.

[7] Equivalent Ability Score based on the British norms.

[8] Equivalent Reading Ages based on the British norms.

Looking first at the comparison of the British and Australian mean scores on each test:

On *Reading Progress Test 3*, the mean score of the Australian Year 3 sample is just under four points higher than the mean score of the British sample at the same year level. The higher mean is to be expected, given that the Australian sample is on average about eight months older than the British sample.

On *Reading Progress Test 4*, the mean score of the Australian Year 4 sample is only slightly higher (less than one point) than the mean score of the British sample at the same year level, even though the age difference is the same as at the Year 3 level. That is to say, the younger British children appear to be catching up with their older Australian counterparts at the same year level.

On *Reading Progress Test 5*, the mean score of the Australian Year 5 sample is about one point lower than the mean score of the British sample at the same year level, while on *Reading Progress Test 6*, the mean score of the Australian Year 6 sample is the same as the mean score of the British sample at the same year level. In other words, by the end of the primary level British children are performing at the same level as Australian children at the same level of schooling, even though they are eight months younger.

Using the British norms as the basis for the interpretation of the Australian scores, this pattern of differences is reflected in the estimated grade-equivalent standardised scores (based on the British norms) for the Australian means on each test – substantially above average (107) on *Reading Progress Test 3*, and about average (100 to 101) on *Reading Progress Tests* 4 to 6.

In terms of equivalent age standardised scores (based on the British norms) the Australian scores are, except at the Year 3 level, substantially lower than would be expected for their age, based on the performance standards of British children at comparable ages.

Converting the mean raw score of the Australian sample on each test to the equivalent ability score based on the British norms indicates that the ability level of each group, relative to that of the British norm group, tends to increase from Year 3 to Year 6.

When the mean scores are converted to reading ages based on the British norms, the reading age of the Year 3 sample is about nine months higher than their chronological age, but at the Year 4 to Year 6 levels the reading age of the Australian sample is about one year below their chronological age. This suggests that by Year 4, reading is more closely related to school level than to chronological age, and the advantage of the Australian group in terms of being older than their British counterparts at the same year level is no longer reflected in their performance level.

These results indicate the need for caution in using reading ages based on the British norms as a basis for interpreting the scores of Australian students on the *Reading Progress Tests*. However, the relatively higher peformance of the Australian Year 3 sample may be due, at least in part, to sampling variation, since the mean standardised scores on the *PAT Reading Comprehension* test indicate that the Year 3 sample is a little above average in terms of reading achievement, with a mean standardised score on the *PAT Reading Comprehension* test of 101.6, while the mean standardised scores of the Year 4 to Year 6 samples on the *PAT Reading Comprehension* test are about average (99.7 to 100.9).

Together with the data on the Stage One *Reading Progress Tests*, these data indicate a shift in the relative performance of Australian and British children through the primary levels of schooling, with Australian children performing better than British children at the same level of schooling up to Year 3, but at about the same level as their younger British counterparts from Year 4 on.

Progression of Scores by Age and Grade

The British norms for the *Reading Progress Tests* are provided in the form of age norms, with separate norm tables for each one month age band. As already indicated, the presentation of norms in this format poses a problem for Australian users, not only because of the mismatch in age at corresponding year levels, but also because of the underlying assumption of a regular increase in score with age, independent of school experience.

In Australia, there are state differences in age of entry to school and in whether or not a pre-Year 1 level is provided prior to entry to Year 1. There are also differences in the period of time spent at the pre-Year 1 level in the case of those states that have a system of continuous enrolment at age five, with some children spending six months or less at the pre-Year 1 level, and other children spending eighteen months or more at the pre-Year 1 level.

In the case of the Stage 1 tests, an examination of the progression of mean scores by age and level of schooling indicated a consistent increase in score according to both age and level of schooling. In view of this finding it was decided to construct the Australian norms for the Stage 1 levels of the *Reading Progress Tests* (the *Literacy Baseline* and *Reading Progress Tests 1 and 2*) on the basis of school level, and to distinguish school level according to both year level and years of school.

In the case of the Stage 2 tests, it was not possible to examine the effect of year of schooling since a different test was administered at each year level. With respect to age, an examination of scores by age within year level indicated some tendency for an increase in score with age at each year level, but this increase was not always consistent across the age range, and was not sufficiently large to justify the construction of age-based norms within a year level. It was therefore decided to construct Australian norms for the *Reading Progress Tests* on the basis of year level.

Table of Norms

Australian norms for *Reading Progress Tests 3 to 6* provide for the conversion of raw scores to standardised scores according to year level. The standardised scores are based on a mean of 100 and a standard deviation of 15, as in the case of the British standardised age scores. In addition to the standardised scores, the table also allows for the conversion of raw scores to percentile ranks and stanine scores.

Standardised scores provide a basis for the interpretation of individual scores relative to the expected performance standards of students at the same year level. They are also useful for comparison of relative performance across different tests or across different areas of the curriculum, and in research studies where standardised measures of performance are required for statistical analysis.

Percentiles are used to describe performance in terms of the percentage of individuals in a particular group who have scores less than or equal to a particular score. For example, a raw score of 20 would have a percentile of 50 if fifty per cent of the group have scores that are less than or equal to 20, a raw score of 30 would have a percentile of 95 if ninety five per cent of the group had scores less than or equal to 30, while a raw score of 10 would have a percentile of 5 if five per cent of the group had scores that were less than or equal to 10. Some users find percentiles useful in reporting information on a student's relative performance.

Stanines provide a means of grouping scores into broader bands that are linked to percentiles and standardised scores. Stanines are commonly used for reporting scores in broad general terms, thus reducing the risk of over-interpretation of small differences. Stanine scores are often grouped into five descriptive categories (well above average, above average, average, below average and well below average) which provide a means of reporting scores in language that is easily understood and interpreted. The stanine levels corresponding to these five descriptive categories, together with information on the expected percentage in each category, and the corresponding percentile level and standard score ranges for each category are shown in Table 3.

Table 3 Descriptive Categories Corresponding to Grouped Stanine Level

Stanine level	Expected %	%ile Range	Standard Score Range	DescriptiveCategory
9	4	96+	127+	Well Above Average
8	7	89-95	119-126	Above
7	12	77-88	112-118	Average
6	17	60-76	104-111	
5	20	40-59	97-102	Average
4	17	23-39	89-96	
3	12	11-22	82-88	Below
2	7	4-10	74-71	Average
1	4	1-3	< 74	Well Below Average

Time of Year Tested

The norms provided are based on data collected at the end of the school year. There is however evidence to indicate that there is little progress in achievement over school holiday periods, and in some cases a decline in performance (see, for example, Tymms, 1999[9]). For this reason the norms based on testing at the end of the school year can be used to obtain an estimate of performance at the beginning of the following year for students at the next year level up. In effect, this means that each test can be used to assess students at either the beginning of the school year or at the end of the school year, provided that the test level appropriate for the previous year level is used at the beginning of the school year (*Reading Progress Test 3* for students at the beginning of Year 4, *Reading Progress Test 4* for students at the beginning of Year 5, etc.). The year-level appropriate test should be used at the end of the school year (*Reading Progress Test 3* at the end of Year 3, *Reading Progress Test 4* at the end of Year 4, etc.).

Using the Norm Table

Table 4 provides for the conversion of raw scores on each test to standardised scores, percentile levels and stanines. For each test, the raw score is listed in the extreme left and right hand columns of the table, while the standardised score (SS) the percentile level (%ile) and the stanine score (St) is listed in the relevant column according to the test administered. Horizontal lines have been inserted to group the scores within each stanine level on each test.

To convert the student's raw score to a standardised score, a percentile level, or a stanine score, first identify the appropriate column according to the test administered. Then locate the raw score in the left or right hand column of the appropriate table, and read across the row to find the corresponding standardised score, percentile level, and stanine score in the appropriate column. These scores should be entered on the class or individual form used for recording student results.

The basic statistics for each norm group are provided at the bottom of each column in Table 4. This information includes the size of the group on which the norms are based (N), the mean raw score and standard deviation of the norm group (Mean, SD), and the average per cent of correct responses on the test. This provides an indication of the difficulty level of the test, with a mean of about 50 per cent of correct responses indicating a test whose difficulty level is appropriate for the group being assessed. Information on the standard error of measurement of the raw score (SEM), and the reliability or internal consistency of the items as measured by the KR-21 formula, is also provided.

[9] Tymms, P. (1999). *Baseline Assessment and Monitoring in Primary Schools: Achievements, Attitudes and Value-Added Indicators.* London: David Fulton Publishers.

Table 4 Reading Progress Tests 3 to 6: Table for Conversion of Raw Scores to Standard Scores

Raw	RPT 3 End Year 3 Beginning Year 4			RPT 4 End Year 4 Beginning Year 5			RPT 5 End Year 5 Beginning Year 6			RPT 6 End Year 6 Beginning Year 7			Raw
Score	SS	%ile	St	SS	%ile	St	SS	%ile	St	SS	%ile	St	Score
51										+			51
50										130	98	9	50
49										126	96	9	49
48							+	+	+	124	95	8	48
47							+	+	+	121	92	8	47
46							130	99	9	118	88	7	46
45				+	99	9	129	97	9	115	85	7	45
44				+	99	9	124	95	8	113	81	7	44
43				+	99	9	121	92	8	112	78	7	43
42				+	99	9	118	89	7	110	75	6	42
41	+	99	9	+	99	9	117	87	7	108	71	6	41
40	+	99	9	130	98	9	115	85	7	105	65	6	40
39	+	99	9	128	97	9	113	81	7	103	59	5	39
38	+	99	9	125	96	8	111	77	6	101	53	5	38
37	130	99	9	122	93	8	109	73	6	100	49	5	37
36	129	97	9	120	91	8	108	70	6	98	45	5	36
35	126	96	8	117	87	7	106	66	6	97	41	5	35
34	122	93	8	115	84	7	105	63	6	95	38	4	34
33	119	90	8	113	81	7	104	60	5	95	35	4	33
32	117	87	7	112	79	7	102	56	5	93	33	4	32
31	115	84	7	110	75	6	101	53	5	92	29	4	31
30	113	80	7	108	71	6	100	50	5	90	25	4	30
29	110	76	6	107	68	6	99	46	5	89	22	3	29
28	108	71	6	105	64	6	97	41	5	87	19	3	28
27	106	66	6	103	59	5	96	38	4	86	17	3	27
26	104	61	6	102	56	5	94	35	4	84	15	3	26
25	102	56	5	101	52	5	93	32	4	84	14	3	25
24	100	51	5	99	48	5	92	29	4	83	12	3	24
23	99	46	5	98	45	5	91	27	4	82	11	3	23
22	97	41	5	97	42	5	90	25	4	82	11	3	22
21	94	35	4	96	39	4	89	23	4	81	10	2	21
20	92	30	4	95	36	4	88	20	3	80	9	2	20
19	90	25	4	93	33	4	86	18	3	80	9	2	19
18	88	21	3	92	30	4	85	15	3	78	7	2	18
17	87	18	3	91	27	4	83	13	3	77	6	2	17
16	85	15	3	90	24	4	82	11	3	76	5	2	16
15	82	12	3	89	22	3	80	9	2	75	5	2	15
14	80	9	2	87	19	3	78	7	2	74	4	2	14
13	79	8	2	85	15	3	76	6	2	73	4	1	13
12	77	6	2	83	12	3	74	4	1	72	3	1	12
11	73	4	1	80	9	2	70	2	1	70	2	1	11
10	70	2	1	78	7	2	-	2	1	-	-	-	10
9	-	1	1	76	5	2	-	1	1	-	-	-	9
8	-	-	-	73	4	1	-	-	-	-	-	-	8
7	-	-	-	71	3	1	-	-	-	-	-	-	7
6	-	-	-	70	1	1	-	-	-	-	-	-	6
N		442			410			381			271		N
Mean		23.7			23.8			29.5			35.4		Mean
SD		7.0			9.2			9.8			9.6		SD
% Correct		57.8			52.9			54.7			65.5		% Correct
SEM		3.0			3.1			3.4			3.5		SEM
KR-21		.81			.89			.88			.89		KR-21

Note: Standardised scores have been calculated to a maximum of 130 and a minimum of 70. Raw score above and below these limits should be recorded as 130 or 70.

OTHER TECHNICAL DATA ON THE READING PROGRESS TESTS, STAGE II

Reading Ages

In the case of the British data, tables are provided for the conversion of raw scores to reading ages. As has already been indicated, these tables are inappropriate as a basis for estimating reading age in Australia because of differences between Australia and the UK in the association between age and year level. There is the further problem that the procedure used for the conversion of raw scores to reading ages is based on the assumption of a linear relationship between test score and chronological age, which is not supported by the data on the relationship between age and score. This leads to serious problems of misinterpretation when reading ages are used as a basis for the interpretation of results.

Reliability

The reliability of a test is a measure of the consistency with which the same results would be obtained by repeated measures using the same instrument. Reliability is conventionally reported as a measure of internal consistency, which is obtained through an analysis of test items.

Measures of internal consistency (KR-21) are reported for each of the norm group samples. The values obtained range from .81 to .82 (see Tables 3 and 4), which are somewhat lower than those reported in the British Manual (.90 to .94). The lower levels may be due to the smaller size of the Australian norm samples, but nevertheless indicate an acceptable level of reliability for the tests when used in an Australian context.

Validity

Validity is a measure of the extent to which a test or instrument measures the behaviours or skills it is intended to measure. While validity can be assessed in various ways, the most common form of validity reported is in terms of correlations with other established measures which are designed to assess the same type of skill. In the case of *Reading Progress Tests 3 to 6*, the British validity data reported are based on test-retest correlations between adjacent tests administered to the same sample of students one year apart (at the end of 1995 and at the end of 1996). These correlations ranged from .69 (between RPT3 and RPT4) to .80 (between RPT5 and RPT6), with these correlations tending to be somewhat higher for the higher level tests than for the lower level tests.

In the case of the Australian data, measures of concurrent validity are available in the form of correlations with the *PAT Reading Comprehension* and *Reading Vocabulary* administered to the same sample of students. These correlations generally range from .7 to .8 (see Table 5), indicating a close correspondence between reading ability as assessed by the *Reading Progress Tests* and reading ability as assessed by the *PAT Reading Comprehension* and *Reading Vocabulary*.

These correlations are also in general consistent with the test-retest correlations reported for the British sample between performance on adjacent levels of the *Reading Progress Tests*.

Table 5 Correlations between the *Reading Progress Tests* and the *PAT Reading Comprehension* and *PAT Reading Vocabulary* tests, by Year Level

	PAT Reading Comprehension				PAT Reading Vocabulary			
PAT Form	Year 3 RPT3	Year 4 RPT4	Year 5 RPT5	Year 6 RPT6	Year 3 RPT3	Year 4 RPT4	Year 5 RPT5	Year 6 RPT6
Form 3-5 A	.73	.76	.74		.72	.72	.78	
Form 3-5 B	.68	.84	.84		.68	.75	.76	
Form 5-7 A			.73	.75			.82	.65
Form 5-7 B			.79	.67			.88	.82

Note: Sample size for each cell ranges from 74 to 189, depending on the combination of forms, except in the case of correlations with the *PAT Reading Vocabulary* test, Form 5-7B (shaded), where sample sizes are lower (15 to 25).