Interpreting ACER Test Results

This document briefly explains the different results provided by the ACER Progressive Achievement Tests (PAT). More detailed information can be found in the relevant teacher manuals.

Understanding the norm tables

The norm table for each of the tests shows the corresponding PAT scale score, percentile rank and stanine score. It also gives the margin of error for each scale score. Note that each PAT assessment has a unique scale. For example, PAT-R Comprehension, PAT-R Vocabulary and PAT Science have unique scales. The exceptions are PATMaths and PATMaths Plus (the online version of PATMaths), which share the same patm scale, and PAT-R Spelling and PAT-SPG Spelling, which share the same pats scale.

Definitions of the results

- The *test raw score* is the number of correct answers on a test.
- The *PAT scale score* is the test raw score converted to the relevant PAT scale. Based on analysis of the data using the Rasch model, this scale enables student achievement and question difficulties to be located on the same scale.
- The *percentile rank* for a particular test raw score shows the percentage of the students tested whose scores fell below that test raw score. It is a measure used to rank candidates in a reference sample, **not a percentage score on the test**.
- A *stanine* provides a coarser ranking than the percentile rank. Stanines divide the total student distribution of abilities into nine categories, with stanine 1 the lowest, stanine 5 the midpoint and stanine 9 the highest.

Explanation of the PAT results

Scale scores

Scale scores are the best measure of students' progress over time. The scale score of an item is a measure of the extent of skills and knowledge required from a student to be successful on the item. A difficult item has a high scale score because it requires more sophisticated skills and richer knowledge to be answered correctly than items lower on the scale. Thus, scale scores take into consideration the level of difficulty of the test items and the level of ability of the students.

For example, a test raw score of 18 on PATMaths Test Booklet 1 is equal to a scale score of 31.4, whereas the same test raw score on PATMaths Test Booklet 2 is equal to a scale score of 40.1. This example shows that relying on test raw scores alone does not give an

accurate picture of a student's ability. Obtaining the same score on both tests could suggest that the two results are equivalent, whereas a comparison of the scale scores shows that the second score is much higher than the first.

Qualities of scale scores

- Scale scores are measures on an interval scale. This means that a difference of 5 in scale scores in the middle of the PAT scale (for example, from 50 to 55) is equivalent to the same difference on any other part of the scale (for example, from 15 to 20 or from 85 to 90).
- Scale scores allow comparison of results on test booklets of varying difficulty.
- Scale scores enable the tracking of students' development in skills as measured by the test from year to year.
- Scale scores provide a common achievement scale for all test booklets giving teachers the flexibility to match test level to student ability and measure growth over time.

Scale score margins of error

All tests have an associated margin of error. Using statistical principles, test developers are able to inform the users of the size of the likely error on any given test score with a great deal of accuracy. These margins of error are often expressed as +/– (plus or minus) a particular value. For example, on the norm table in the *PATMaths Teacher Manual* (page 65), a student with a test raw score of 16 on PATMaths Test Booklet 2 would have a corresponding scale score of 37.6. The error margin for this score is 3.5, so the student's real score is likely to fall between 34.1 and 41.1 (37.6 plus and minus 3.5).

When you are interpreting results it is important to pay attention to the inherent error in all measurement. Small differences in scores should not be given more importance than they deserve. Error margins are larger for very high and very low scores.

The norm tables in the teacher manual for each test provide information on the error margins associated with each test score.

Percentile ranks

Percentile ranks provide a simple means of indicating the rank order and position of a student's result in relation to a norm-reference sample.

For example, a student's percentile rank of 45 means that 45 per cent of the reference sample achieved a test score lower than or equal to the test score obtained by the student. A student's percentile rank of 96 means that 96 per cent of the reference sample achieved a test score lower than or equal to the test score obtained by the student.

Most PAT norm-reference samples consist of Australian students from all states and territories and from all school systems – government, Catholic and independent. A random, stratified sampling frame is used. This allows your school to compare your students' performance against a large Australian reference sample. By comparing your school's performance against a reference sample, it is possible to see if your students' performance on the test is stronger or weaker than the reference sample.

Local norms can be provided through ACER Test Scoring Services if a school has more than 100 students, in a particular year level, sit the same test at the one time. This means that in addition to national reference group rankings, students can be ranked on a school basis.

Stanine scores

Stanine scores are derived from percentile ranks. Percentile ranks are divided into nine categories called stanines (short for 'standard nine') and the digits '1' to '9' are used as category labels.

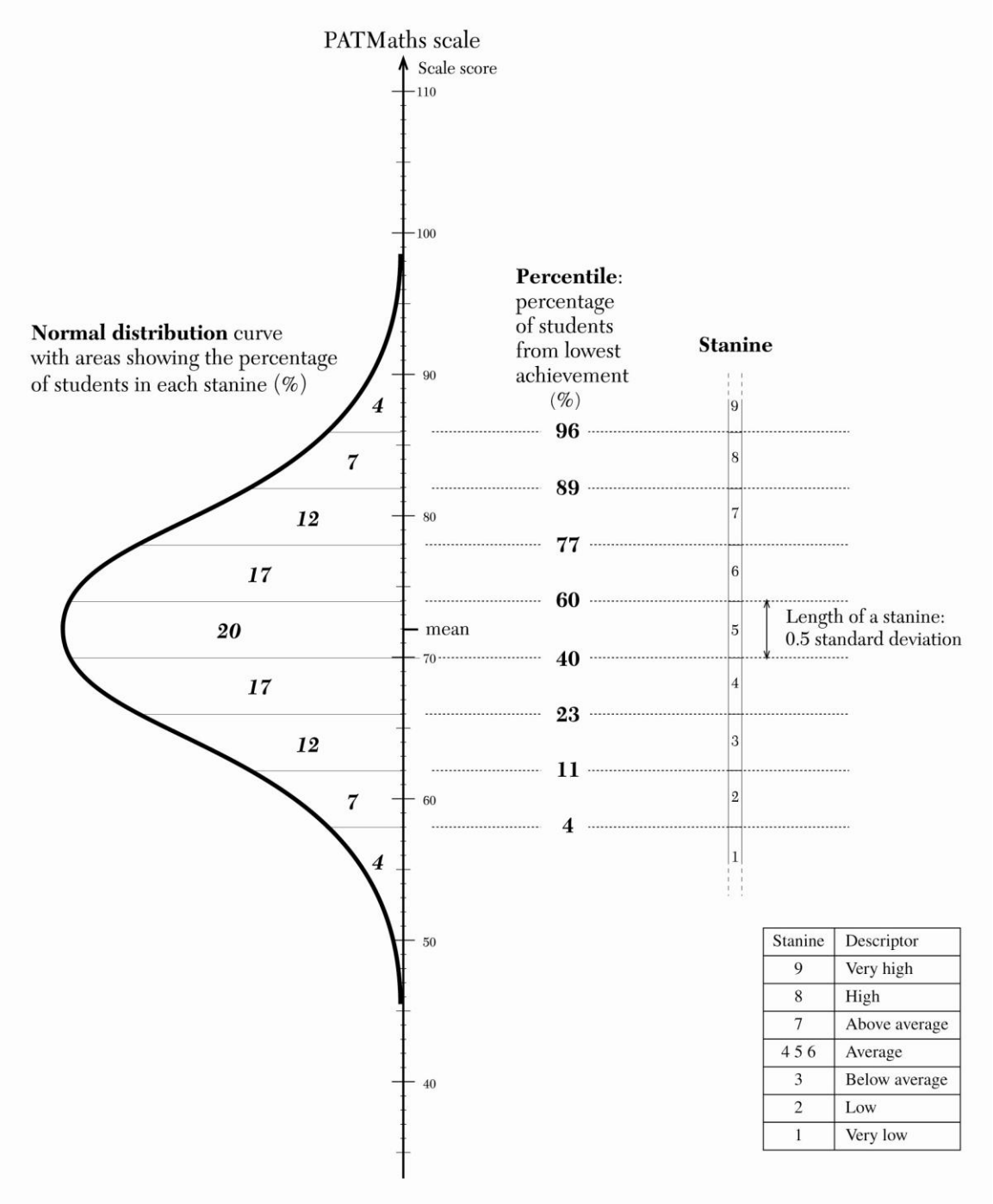
Stanine scores are particularly useful for grouping students; however, it is recommended that only differences of two or more stanines should be regarded as indicating a real difference in performance.

Relationship between percentile ranks and stanine scores

The bell-shaped curve on the next page shows how the stanine is derived. The underlying basis for obtaining stanines is that a normal distribution is divided into nine intervals, each of which has a width of one half of a standard deviation excluding the first and last stanine. The mean lies approximately in the centre of the fifth interval.

The table below shows the relationship between stanines and percentile ranks.

Normative description of student achievement	Stanine	Corresponding percentile rank	Percentage of students
Very high	9	96 and above	4
High	8	89–96	7
Above average	7	77–89	12
Average	6	60–77	17
Average	5	40–60	20
Average	4	23–40	17
Below average	3	11–23	12
Low	2	4–11	7
Very low	1	0–4	4



Stanine	Descriptor
9	Very high
8	High
7	Above average
4 5 6	Average
3	Below average
2	Low
1	Very low

Technical information

All ACER PAT teacher manuals provide detailed technical information. Refer to the appropriate chapter in each manual for information about construction of each PAT scale, validity, reliability and characteristics of the norm-reference groups.