

## **Interpreting Scores**

### **Scale Scores**

Scaled scores take into consideration both the level of difficulty of the test items and the level of ability of students. One of the strengths of scaled scores is that they allow a comparison between test forms of varying difficulty.

For example, a raw score of 18 on PAT Maths Form 1 is equal to a scale score of 31.4, whereas the same raw score on PAT Maths Form 2 is equal to a scale score of 40.1. It can be seen from this example that relying on the raw scores alone does not give an accurate picture of a student's ability. Obtaining the same raw score on both tests could suggest that the two results are equivalent whereas a comparison of the scaled scores shows that the second score is much higher than the first.

Provision of scaled score also enables the tracking of students' development between school years. For example with the Progressive Achievement Tests in Mathematics (PAT Maths) 3rd Edition the scaled score provides a common achievement scale for all eight tests giving teachers the flexibility to match test level to student ability and measure growth over time.

### **Percentile Ranks**

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

For example, a student with a percentile rank of 45 has a score that is higher than the score obtained by 45 per cent of the reference group students, whilst a student with a percentile rank of 96 has a score that is higher than the score obtained by 96 per cent of the reference group. The higher the percentile rank, the better the score.

Most of the tests scored through Test Scoring Services provide percentile ranks against a normed population. This allows the school to compare their students' performance against a larger reference group. By comparing the school's performance against the reference group it is possible to see if their students' performance on the test is stronger or weaker than the reference group.

Local norms can also be provided if a school has more than 100 students, at a particular year level, sit the test at any time. This means that the students are ranked against themselves rather than the larger reference population.

### **Stanines**

Stanines are derived from percentile ranks. Percentile ranks are divided into 9 sub divisions – called stanines (short for 'Standard Nine'), and the digits 1 to 9 are used.

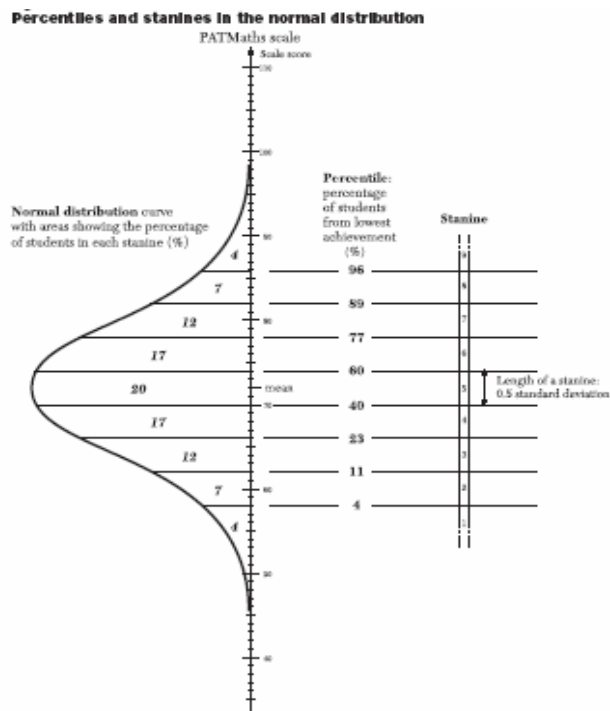
Stanines are particularly useful for grouping students, however it is recommended that only differences of two or more stanines should be regarded as indicating a real difference in performance.

The following table provides a descriptor of students' performance against the scored stanine.

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

### Relationship between Percentile Ranks and Stanines

The normal curve below 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. The mean lies approximately in the centre of the fifth interval.



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

Description of Performance	Stanine	Corresponding percentile rank	Percentage of cases	Number of students in a class of 25
Very high	9	96 and above	4	•
High	8	89-96	7	• •
Above average	7	77-89	12	• • •
	6	60-77	17	• • • •
Average	5	40-60	20	• • • • •
	4	23-40	17	• • • •
Below average	3	11-23	12	• • •
Low	2	4-11	7	• •
Very low	1	0-4	4	•

Note: Each • represents one student.

### Error in Scores

When 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 manuals associated with each test provide information on the error margins associated with the test scores.

For example, in the PAT-R Third Edition Manual on page 61, it shows that a student with a raw score of 16 on PAT-R Test Form 2 would have a corresponding scale score of 39.0. The error margin for this score is 3.6. That is we can be reasonably confident that the score will be 39.0 plus or minus 3.6, that is between 35.4 and 42.6.