

CAT-4 provides reports at both the student and class levels. Interpreting these reports for the multiple-choice components requires some understanding of the difference between norm-referenced scores and criterion-referenced scores. A *norm-referenced* score reveals how a student did on a test relative to other same-grade students. A criterion-referenced score reveals how a student did on part of a test relative to pre-defined performance criteria.

The In-Service Folder provides an example of a student-level report called the **Student Test Record**. The left column contains information about the student and the tests. The name of the student and teacher are presented first. The student's grade is 6.1, meaning that the student wrote the tests after completing one month of grade 6. In other words, the student wrote the tests in October. The student wrote the Level 15 tests, which cover the curriculum up to the end of grade 5. The test date and school and district/division names are also provided.

Norm-Referenced Scores

One way to look at how well a student did on a test compared to others is to report the percent of students who got a lower score than that student's score. This type of score is called a **percentile rank**. When the student's score is compared to the scores of other same-grade students across Canada, it is called a *national* percentile rank. When the student's score is compared to the scores of other same-grade students in one region of Canada (e.g., Western Canada), it is called a *regional* percentile rank. When the score is compared to the scores of other same-grade students in the district or division, it is called a *local* percentile rank. These percentile ranks are shown, near the top part of the Student Test Record, under "Norm-Referenced Scores." For example, the student's National Percentile for Reading is 42, meaning that the student's Reading score is higher than 42% of the scores attained by the national sample of same-grade students. Percentile ranks range between 1 and 99.

The 99 possible percentile rank scores can be grouped into larger categories called **stanines**. Stanines range between 1 and 9, where 9 is the highest possible score. Table 1 shows how percentile ranks are grouped into stanines.

Table 1

| STANINE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|------------------|-----|------|-------|-------|-------|-------|-------|-------|-------|
| PERCENTILE RANKS | 1–4 | 5–10 | 11–22 | 23–40 | 41–59 | 60–77 | 78–89 | 90–95 | 96–99 |

As is the case for percentile ranks, there are *national*, *regional* and *local* stanines. These stanines are shown in the Student Test Record under "Norm-Referenced Scores." For example, the student's National Stanine for Reading is 5.

Percentile Ranks and stanines are also reported for total scores. A total score is a weighted combination of two or more test scores. For Levels 14 to 19, Total Reading is the weighted average of the Reading and Vocabulary scores, where Reading is given more weight (75%). Total Writing Conventions is the weighted average of the Writing Conventions and Spelling scores, where Writing Conventions is given more weight (80%). Total Math is the weighted average of the Mathematics and Computation scores, where Mathematics is given more weight (67%). Total Battery is the weighted average of Total Reading (40%), Total Writing Conventions (20%) and Total Math (40%).

In the sample report, the graph in the top right corner shows the student's national percentile ranks (the scale at the top of the graph) and the corresponding national stanines (the scale at the bottom of the graph). The wide shaded column shows the average score range. The area to the left of the shaded column shows the below average range, and the area to the right of the shaded column shows the above average range. The big dots on the graph show the student's percentile rank/stanine for each test. The lines on either side of a dot show an estimate of how much error can enter into the estimation of a student's ability from this single test score. The idea is to identify which region the line falls in mostly. For example, most of the line for Reading falls in the shaded Average region, so the student's reading ability is average compared to other same-grade students across Canada. Most of the line for Mathematics falls in the Above Average area (to the right of the shaded region) so the student's mathematics ability is above average compared to other same-grade students across Canada.

Another way to compare a student's ability to the abilities of other students is to estimate the student's *grade equivalent*. The sample report shows the grade equivalent 5.3 for Reading. This means the student's score for Reading indicates a level of achievement that is equivalent, on average, to the achievement of students who had completed 3 months of grade 5. Since the student has completed one month of grade 6, the student's grade equivalent for Reading indicates that his reading ability is below grade level. The lowest grade equivalent reported for the Level 15 test is 3.9, so the score 3.9– would have been reported had the student

achieved a grade equivalent lower than 3.9. The highest grade equivalent reported for the Level 15 test is 7.9, so the score 7.9+ is reported whenever the student achieved a grade equivalent higher than 7.9. A grade equivalent of 7.9 would reflect a level of achievement equivalent, on average, to students who had completed 9 months of grade 7.

A caution is in order in the use of grade equivalents. Grade equivalents should be used only to estimate whether a student’s achievement is below, at, or above grade level compared to other students in Canada. An unusually high or low grade equivalent may indicate a need for enrichment or intervention. However, a grade equivalent should never be used alone to make high-stakes decisions, such as whether or not to place a student in a gifted program. (If a grade 5 student is administered a grade 5 test and gets a grade equivalent of 7.5, it doesn’t mean that the student can do grade 7 work. The student hasn’t even been exposed to the grade 6 curriculum yet, let alone the grade 7 curriculum.)

Table 2 provides the definitions and characteristics of norm-referenced scores for easy reference.

Table 2 Norm-Referenced Information

| | PERCENTILE | STANINE | GRADE EQUIVALENT |
|-------------------------|--|--|---|
| WHAT IT IS | How student’s score compares to the scores of other students in the same reference group | Each percentile rank is placed in one of nine categories | Estimate of student’s grade-level functioning on a given test |
| REFERENCE GROUPS | National (NP) Regional (RP)* District (LP) | National (NS) Regional (RS)* District (LS) | National (GE) |
| POSSIBLE SCORES | (low) 1 to 99 (high) | (low) 1 to 9 (high) | 0.1 to 12.9 |
| WHY IT IS USEFUL | Shows precisely how a student is performing in comparison with others at the same grade | Provides quick and simple information | Shows a student’s progress from year to year |
| WHY IT CAN BE CONFUSING | It is so precise that variations are bound to occur, partly from chance | Is only sensitive to very large changes in student performance | Grade equivalent does not indicate that the student should be moved up or down to another grade |

* Regional norms are available for Western Canada, Ontario and Atlantic Canada

Criterion-Referenced Scores

Criterion-referenced scores are shown in the bottom part of the Student Test Record. For each test objective, the student’s performance is classified as Low, Competent or Proficient. Competent means “meeting end-of-grade expectations” and Proficient means “exceeding end-of-grade expectations.” Low means “not meeting end-of-grade expectations.” (In this case, the student wrote the Level 15 tests at the beginning of grade 6, so the expectations are end-of-grade 5 expectations.) The student’s classification for each objective is determined by the number of questions the student answered correctly for the test questions under that objective. In the sample report, the student attained a Competent classification for Literary Text (under Reading). This means that the student correctly answered enough of the Literary Text questions to be classified Competent, but did not correctly answer enough Literary Text questions to be classified Proficient.

A caution is in order when looking at criterion-referenced scores for *individual students*. CAT-4 objectives are usually represented by 12 test questions. Although this is a greater number of questions than the number most large-scale assessments use for reporting sub-test scores (e.g., objective scores), it is not enough to provide highly reliable scores at the *individual student level* (but can provide very reliable

information at the class or school levels). For this reason, it is important to look for patterns in a student's criterion-referenced scores rather than to over-interpret a single criterion-referenced score in isolation.

When looking at the Student Test Record, one may sometimes find an apparent “discrepancy” between the norm-referenced scores and criterion-referenced scores for the same test. In this example, the student's National Percentile for Writing Conventions is 90, and the National Percentile Range falls mostly in the Above Average region. However, the student attained a Proficient score in only one of the three objectives for Writing Conventions (Sentence Structure). It is important, however, to remember that a Percentile Rank indicates how the student did *compared to other students*, whereas criterion-referenced scores indicate how a student did *compared to pre-defined performance criteria* (independently of how other students did). The current example indicates that the student is above average compared to other students and is meeting the performance standards for two of the objectives and exceeding the performance standard for one objective. This simply means that this student has done a better job at meeting/exceeding the performance criteria than most other students in Canada.

The In-Service Folder provides an example of a class-level report called the **Objective Competency Report**. This report provides summary criterion-referenced score information for a class. The top left corner of the table shows that the class consists of 22 students. Below this information is a list of the test objectives. The column to the right of the list shows the number of students in the class who scored Low, Competent and Proficient for each objective. Scanning the column for Low can help the teacher identify the objectives for which students in her class are having the most difficulty. The columns in the main part of the table show the information for each student (listed alphabetically by last name). Scanning down a column for a student gives the teacher an idea of the student's strengths and needs. Scanning across the row of students for a particular objective gives the teacher an idea of which students had difficulty with that objective.

The information in the Student Test Record and Objective Competency Report can be used to help plan instruction. The matching of the test questions to the provincial curricula can be found at www.canadiantestcentre.com. For more information, refer to “Linking Assessment to Instruction at the Student and Class Levels.” For additional support of the teacher's use of test results, the CAT-4 *Teacher Resource Manual* provides instructional activities organized by the CAT-4 test objectives. The *Teacher Resource Manual* can be purchased at www.canadiantestcentre.com or by calling 1 800 668 1006.