

## 8.0 GLOSSARY

Terms in **boldface type** appear within the glossary as a separate entry.

**Achievement Levels**—Five categories of achievement that represent the success students demonstrate with the *Sunshine State Standards* content assessed on the FCAT. Achievement Levels are established using the input of classroom teachers, curriculum specialists, education administrators, and other interested citizens. These professionals helped the DOE identify the score ranges for each Achievement Level. The Achievement Levels are helpful in interpreting what a student’s **scale score** represents.

**Anchor Items**—A common set of **items** on tests administered in two different years used to develop comparable **scale scores**. The student performance on the anchor items is the source for the data used in the statistical **equating** procedures.

**Anchor Papers**—Student responses that demonstrate typical performance for each score point in the **rubric**. As they score each student’s response to an **item**, scorers compare the response to the anchor papers to determine the number of points that student has earned. Also called **rangefinders**.

**Backreading**—Method used to ensure adherence to scoring guidelines for **performance task items** and essays. Scoring officials evaluate appropriateness of scores assigned by scorers.

**Benchmark**—A statement within the *Sunshine State Standards* that describes what students at a certain grade level should know and be able to do. More detailed than a **strand** or **standard**.

**Bias**—Advantage or disadvantage conferred upon groups of students because of certain personal characteristics (such as gender, race, ethnicity, religion, socioeconomic status, disability, or geographic region) unrelated to an understanding of the content.

**Calibration Sample**—Carefully selected group of students representative of all students statewide whose response data are used to generate **Item Response Theory (IRT)** parameters used in **operational** test scoring.

**Census Test**—The assessment of all eligible students at a particular grade level in a specific **content area**. This term is used to distinguish an assessment of all students from an assessment of only a sample of students.

**Cloze**—Text with blanks inserted where a word or words need to be added. After reading the cloze sample, students choose the answer that correctly completes the sentence. FCAT Writing+ cloze samples contain high-interest material in a relatively short format that can be more literary or technical in nature than the text in the other sample types. Cloze samples are not presented as representative of student-generated work. On a test form, each cloze sample contains three to four numbered blanks used to measure the student’s knowledge of spelling or usage conventions.

**Cluster (content cluster)**—A grouping of related **benchmarks** from the *Sunshine State Standards*. Clusters are currently used to summarize and report achievement for FCAT Reading and FCAT Science and, beginning in 2006, will be used to summarize and report achievement for FCAT Writing+.

**Cognitive Complexity**—System used to classify FCAT **items** according to the complexity of the steps and processes they require students to use.

**Content Area**—The information or skills contained in an area of study. The content areas (or subject areas) assessed on the FCAT are reading, mathematics, science, and writing.

**Content Area Subscores (content cluster scores, content area scores)**—The number of points earned by a student in each **cluster** or **strand** of the *Sunshine State Standards* portion of the FCAT. Content subscores are reported for **clusters** in FCAT Reading and FCAT Science. In FCAT Mathematics, content subscores are reported for **strands**. Beginning in 2006, FCAT Writing+ content subscores will be reported for **clusters** and for the essay. Computed before **IRT** processing and **equating**.

**Content-Sampled Benchmarks**—**Benchmarks** assessed periodically (as opposed to annually) by FCAT Science.

**Criterion-Referenced Test (CRT)**—An assessment where an individual’s performance is compared to a specific learning objective or performance standard and not to the performance of other students. Criterion-referenced tests show how well students performed on specific goals or **standards** rather than just telling how their performance compares to a norm group of students nationally or locally. The FCAT, a CRT, is based on the *Sunshine State Standards* and measures student progress toward meeting these **standards**.

**Cut Point Scores**—FCAT **scale scores** or FCAT **developmental scores** that mark the boundaries between different **Achievement Levels**.

**Developmental Scale Score (DSS)**—A type of **scale score** used to determine a student’s annual progress from grade to grade. Calculated by converting a student’s **scale score** (100–500) to a scale from 0 to about 3000 that is used for Grades 3–11.

**Equating**—A process used to place **IRT**-processed scores on the FCAT scale of 100 to 500 and to ensure that the resulting scores are comparable to those of previous years. Students are tested in two different years with tests that have a common set of **items** called **anchor items** as well as different **items**. The **anchor items** and how students perform on them from year to year are used in the statistical equating procedures. Equating scores ensures that the same standard of achievement is used each year so the progress of students and schools can be evaluated fairly, i.e., Grade 8 scores in 2004 are comparable to Grade 8 scores in 1998.

**Exceptional Student Education (ESE)**—Special educational services that are provided to eligible students, e.g., visually impaired or hearing impaired. These services are required by federal law and are provided to Florida students according to the State Board of Education Rule 6A-6.0331, FAC. Also known as Students With Disabilities (SWD).

**Expository Writing**—Writing that gives information, explains why or how, clarifies a process, or defines a concept. In FCAT Writing+, students in Grades 4, 8, and 10 are assigned **prompts** intended to elicit expository writing.

**Extended-Response Item (ER)**—See **Performance Tasks**.

**FCAT Score**—For FCAT Reading and FCAT Mathematics, the **Developmental Scale Score**. For FCAT Science, the **scale score**.

**Field-Test Item**—**Item** included on the FCAT for **item** development purposes only. Student response data are reviewed to determine whether a field-test **item** would be a useful **operational item**. Does *not* count toward student scores.

**Gridded-Response Item (GR)**—Test **items** that require students to solve a problem for which the answer is numerical. Answers must be written and bubbled into a number grid. The gridded-response **item** format is used in FCAT Mathematics (Grades 5–10) and FCAT Science (Grades 8 and 11).

**Holistic Scoring**—A method of scoring written work that considers the overall quality of the entire work. Scores are assigned to student work using a pre-defined **rubric**.

**Individual Education Plan (IEP)**—Describes special education services provided as part of **Exceptional Student Education**. Also specifies the testing accommodations a student needs for classroom instruction and assessments.

**Item**—Any test question or task for which a separate score is awarded.

**Item Bank**—Database of **field-test** and **operational items**. **Items** are selected from it each year to construct the FCAT.

**Item Response Theory (IRT)**—Statistical model for student responses to test **items**. Based on the idea that the likelihood of student success on an **item** is the result of the student’s true level of ability and three characteristics of the **item**: ability of the **item** to differentiate between students at different **Achievement Levels** (the *a*-parameter), difficulty of the **item** (the *b*-parameter), and the effectiveness of guessing (the *c*-parameter, for **multiple-choice items** only). Used solely in FCAT **item** and test development and as the basis of generating **scale scores**.

**Limited English Proficient (LEP)**—Special education services for non-native speakers of English. LEP students, also known as English Language Learners (ELL), are permitted testing accommodations when taking the FCAT.

**Linking**—Method used to create **developmental scale score**. A small sample of identical **items** are given to students in adjacent grades.

**Mode of Writing**—Characteristics of written work that reveal the purpose of the writing. The essay portion of FCAT Writing+ assesses three modes of writing: **narrative**, **expository**, and **persuasive**.

**Multiple-Choice Items (MC)**—**Items** that present students with several options from which to choose. FCAT Reading, Mathematics, Science, and Writing+ multiple-choice **items** have four choices, only one of which is correct. Writing+ has some three-option multiple-choice **items**.

**Narrative Writing**—Writing that tells a story based on a real or imagined event. In FCAT Writing+, only students in Grade 4 are assigned a **prompt** intended to result in narrative writing.

**Norm-Referenced Test (NRT)**—A test designed to compare the performance of one group of students to a national sample of students, known as the “norm” group. The NRT portion of the FCAT includes both the Reading Comprehension and Mathematics Problem Solving subtests from the *Stanford 10* test published by Harcourt Assessment, Inc.

**Operational Items**—**Items** that count toward a student’s score. Most **items** on the FCAT are operational **items**.

**Pattern Scoring**—A method of calculating a test score based on comparison of students’ overall patterns of success on **items**. Pattern scoring shows inconsistencies in student responses (i.e., lack of success on an **item** with the same level of difficulty as other **items** with which the student had success).

**Performance Tasks**—**Items** that require students to provide either a short or extended written response. For example, short-response (SR) tasks may ask students to describe a character in a story, write a mathematical equation, or explain a scientific concept. Examples of extended-response (ER) tasks may include comparing two characters, constructing a graph, or describing the steps in an experiment.

**Persuasive Writing**—Writing that attempts to convince the reader that an opinion is valid or that the reader should take a specific action. In FCAT Writing+, students in Grades 8 and 10 are assigned **prompts** intended to result in persuasive writing.

**Pilot Test**—An assessment of a sample of students for the purpose of gaining general information about students' reactions to test **items**. Statistical analysis is not the focus of this initial tryout of **items**.

**Plan-Based Items**—A writing plan provides a prewriting structure and is based on a topic that is within the purview of students at the specified grade level. Possible graphic organizers may include charts, webs, diagrams, and outlines. In FCAT Writing+ tests, students answer questions about strengths and weaknesses of the writing plan.

**Prompt**—The topic a student is given on which to write an essay in FCAT Writing+. The **prompt** has two parts: the *writing situation* (presents and clarifies the topic) and the *directions for writing* (guides the student to think about the topic and suggests an approach that may help the student begin writing).

**Rangefinders**—Student responses to **prompts** (FCAT Writing+) or **performance tasks** (FCAT Reading, Mathematics, or Science) used to illustrate score points on the **rubric**. Rangefinding is the process of identifying these student responses. Also called **anchor papers**.

**Raw Score**—A score that reports the number of points a student earned on each test **item**, **cluster/strand**, or the entire test. Students earn one raw score point for each correctly answered **multiple-choice item** and **gridded-response item**, and up to four raw score points on **performance tasks**. Raw scores are reported as **content subscores**.

**Released Item**—A test question that has been released to the general public.

**Reliability**—Desired characteristic of a test. Achieved when measurement error is minimized and the test score is close to the **true score**.

**Retake**—Alternate Grade 10 reading or mathematics test given to those who do not achieve the passing score required for high school graduation.

**Rubric**—Scoring guidelines or criteria used to evaluate all FCAT **performance tasks** and essays. Describes what characterizes responses at each score point.

**Sample-Based Items**—A writing sample is an example of draft writing. Writing samples may be draft stories, reports, or articles that contain some mistakes. FCAT Writing+ **items** based on writing samples ask about the strengths and weaknesses of the sample.

**Scale Score**—Score used to report student results for the entire test in FCAT Reading, Mathematics, and Science. Scale scores on the FCAT range from 100 to 500 at each grade level. The scale score is the result of **IRT** processing and **equating**.

**Section 504**—Special classification of students as defined in Section 504 of the Rehabilitation Act of 1973. Testing accommodations are permitted for students who meet the Section 504 criteria.

**Short-Response Item (SR)**—See **Performance Tasks**.

**Stand-Alone Items**—Provide a succinct context for measuring the student’s knowledge of the conventions of capitalization, punctuation, and sentence structure to address the breadth of the FCAT Writing+ editing **benchmark**.

**Standard**—In the *Sunshine State Standards*, a statement of what students should know and be able to do. More specific than a **strand** and not as specific as a **benchmark**.

**Standard Error of Measurement (SEM)**—A whole-test **reliability** indicator that is calculated using data from the entire tested population. For example, if a student were to take the same test over and over (without additional learning between the tests or without remembering any of the questions from the previous tests), the difference in the resulting test scores is called the standard error of measurement.

**Strands**—The broad divisions of **content areas** in the *Sunshine State Standards*. For example, in the Language Arts *Sunshine State Standards*, there are seven **strands** (Reading, Writing, Listening, Viewing, Speaking, Language, and Literature).

**Sunshine State Standards (SSS)**—Florida’s curriculum framework that provides guidelines for what students should know and be able to do in each subject at each grade. Describes learning expectations at increasingly more detailed levels: **strands**, **standards**, and **benchmarks**. The purpose of the FCAT is to measure the *Sunshine State Standards* **benchmarks**. All FCAT **items** are based on specific **benchmarks**.

**Test Form**—A unique set of **items** consisting of a common core of **operational items** and a smaller number of either **field-test** or **anchor items**. FCAT Reading, Mathematics, and Science all use multiple test forms. Students with different test forms face exactly the same **operational items** but different **field-test** or **anchor items**.

**True Score**—FCAT seeks to measure a student’s “true” achievement or true score on the content assessed. By definition, a student’s test score is composed of two parts—the true score and the **standard error of measurement** associated with the test.

**Validity**—Desired characteristic of a test. Achieved when the test actually measures what it is intended to measure.