FLORIDA DEPARTMENT OF
EDUCATION
—fldoe.org—

Florida's Statewide Grades 3–10 and End-of-Course (EOC) Computer-Adaptive Tests (CAT) Frequently Asked Questions (FAQ)

The purpose of this document is to provide information about the adaptive functionality of Florida's statewide assessments. For the assessments described in this document, Cambium Assessment is the testing vendor that provides the computer-based testing platform and adaptive algorithm.

1. **What is a Computer-Adaptive Test (CAT)?**

   CATs are administered online. In a CAT, each test question is selected by a computer algorithm that meets the published test blueprint addressing the content standards and is based on the student's performance on previously administered items. A CAT is tailored to the student's ability based on whether the student responded successfully to the preceding item or sets of items. Each time a student answers a question, his or her response helps determine the next question or set of questions that will be presented to the student.

2. **Which tests are CAT?**

   The following assessments, launched in the 2022–23 school year, are CAT:

   - Florida Assessment of Student Thinking (FAST) VPK–2 Early Literacy, Reading, and Mathematics (not addressed in this FAQ)
     - For more information about K–2 assessments, please see the 2023–24 FAST Kindergarten–Grade 2 Fact Sheet.
   - FAST Grades 3–10 English Language Arts (ELA) Reading
   - FAST Grades 3–8 Mathematics
   - Benchmarks for Excellent Student Thinking (B.E.S.T.) Algebra 1 and Geometry EOC assessments

   Beginning in the Spring 2024 administration, the following assessments will be CAT:

   - Grades 5 and 8 Science
   - Biology 1, Civics, and U.S. History EOC assessments

3. **What is the benefit of a CAT?**

   There are several benefits to CAT:

   - The test attempt is tailored to the ability level of a student, so the precision of resulting test scores is increased, across all ability levels.
   - The tests are more precise because the test is adapting dynamically to the students' ability level, so fewer items can be delivered to students than a fixed-form test.
   - The tests are more secure since students encounter different items throughout the test.
   - Results are delivered within 24 hours after the student takes the test.

4. **How does the CAT algorithm work?**

   The algorithm selects test items from a common item bank that contains test items aligned to the Florida academic content standards. The item selection is driven by two things: the test blueprint (described in #6 below) and student performance. Item selection is primarily driven by the test blueprint, ensuring that students are administered tests that represent the breadth and depth of the standards. As mentioned in previous questions, how a student responds to an item informs the algorithm as to whether it selects more or less challenging items to appear next.

The algorithm functions slightly differently depending on the test being administered:

- **Mathematics:** adaptive at the item level.

  - Based on a student's response to an item and blueprint requirements, the algorithm will select the next appropriate item.

- **Reading:** adaptive at the passage/passage set level.

  - Items on a Reading test are attached to a specific passage and are, therefore, referred to as a "passage set." Therefore, based on a student's responses to a passage set (rather than an individual item) and blueprint requirements, the algorithm will select the next passage/passage set.
  - Once the algorithm selects the next set, it is locked in (regardless of whether students go back and change answers within the current set).

- **Science and Social Studies (spring 2024):** adaptive at the test blueprint level.

  - Each item a student receives will be randomly selected to meet blueprint requirements and cover a range of item difficulty levels. Beginning in Spring 2025 and beyond, the tests will be fully adaptive at the item level, similar to Mathematics tests.

5. **On FAST tests, do the items selected by the algorithm go above and below the student's current grade level?**

No; the item bank for each test only contains items within the student's current grade level.

6. **What is a test blueprint?**

The ELA and Mathematics blueprints show the percentage of items in each reporting category students will encounter during each progress monitoring (PM) window. Students will see the full scope of grade/course-level content that falls within the percentage of the overall test length. Blueprint coverage ensures that each student sees a certain percentage of items from each reporting category; they will not necessarily see an item for every benchmark.

7. **How does the algorithm select the first item for each student?**

The first time a student takes a test in a given subject, the algorithm will choose a mid-level difficulty item for the student. If the system has prior data for a student (if they took a test in the same subject within the past year), the algorithm will pick up where the student left off in terms of item difficulty. Exceptions to this are grade 3 ELA Reading and Mathematics, where students are transitioning from Star assessments to FAST, and the EOC assessments, which do not track data between attempts for students.

8. **If a student starts with a low-level item, does it affect his or her ability to achieve a high scale score?**

No. The test is sufficiently long enough for all students to be presented with test content that allows them to demonstrate their true ability. Although a student may start with easier questions based on performance on a prior PM, if he or she has increased knowledge, skills, and abilities, the test will adapt appropriately to allow them to move up one or more achievement levels.

9. **How is the difficulty level determined for each item?**

Data analyses conducted after field testing and after operational testing generate statistics for each item. These statistics inform the degree to which the item differentiates between students of different abilities, the difficulty of the item, and the likelihood of success by guessing. These item statistics, along

with others, determine an item's level of difficulty. For instance, an item with a low percentage of students correctly answering the items is likely to have a higher difficulty level.

The algorithm selects items first on the published test blueprint. Then, from the items that measure that standard, it selects the most appropriate item based on item difficulty and the student's performance on previous items.

10. If students enter an incorrect response to an item, will the algorithm select a lower-level item next even if it is from a different reporting category?

The algorithm discriminates by overall performance, not reporting category. It will always select the next item based on the blueprint requirements and the estimate of the student's ability. This estimate becomes more reliable as the student progresses through the test.

11. Will students all receive the same number of items?

On most assessments, students will receive the same number of items. On FAST PM3, students will see field test items on their assessment. Depending on the number of field test items a student receives, their test may be a few items longer or shorter.

12. Can students skip items?

No. Students must provide an answer for each item to move on to the next item. Students should always provide their best answer for each item as they encounter it.

After a student answers each item, the adaptive algorithm updates its interim ability estimate (or interim skill level). Based on that information, plus the blueprint coverage requirements, the algorithm searches for the next best item to be administered. This cycle continues until the end of testing, and it results in a tailored test form with more engaging questions. The final ability estimate (or scale score) is derived from a student's final response set across all items seen during the test. A student must answer each question so that the adaptive algorithm can search for the next best item at each step of testing. In other words, students are not allowed to skip questions.

13. Can students return to an item and change the answer?

Yes. Students may return to items, and they are able to change the answer they provided if they choose to do so.

14. How does going back and changing an answer affect the algorithm?

Returning to an item and changing an answer does not affect the algorithm since the subsequent items have already been selected and delivered. However, changing an answer could affect a student's overall score (if it was changed from right to wrong or wrong to right).

If a student is unsure about an earlier item and goes back to change a previously provided response(s), the algorithm is not going to pull a new set of questions, but the ability estimate (or the scale score) will be derived though a statistical process, based on students' final response set.

15. What is the "Mark for Review" tool and does using it affect the test algorithm?

The Mark for Review tool is provided on all Florida computer-based tests, and it allows students to flag an item they would like to review later after providing their best answer. Students select "Mark for Review" from a drop-down menu on the item, and a flag is placed on that item on the item review screen to remind the student that they may want to return to that item later. Using this tool does not affect the algorithm in any way.

16. **Are there specific test strategies that students should use when taking a FAST or B.E.S.T. CAT?**

    Because the test is computer-adaptive, students should do their best on each item as they progress through the test. Students are not able to preview all items in the test because they have to provide a response to each item (Mathematics) or passage set (Reading) before proceeding.

17. **How is the overall scale score determined?**

    Florida's statewide assessments are scored using a method called "pattern scoring." This means that the pattern of answers provided by a student is analyzed in combination with the item statistics. In other words, information about the pattern of answers (which questions were answered) and the statistical qualities of test items (difficulty level) are evaluated together to determine the scoring weights for each item and the likelihood of an individual student's score.

    As a result of this method of scoring, students with the same raw score may have similar, but not necessarily identical, scale scores. Students who were successful with the more challenging content in the item bank will have a higher scale score.