Item Response Theory and CATs in FAST™

Adaptive Math (aMath) and Adaptive Reading (aReading) use more precise measurement models than most other computer adaptive tests (CATs). The CATs in FAST™ consider three important characteristics of each assessment item, whereas others rely on only one. The models that drive CATs in FAST™ are more precise and, thus, aReading and aMath can be more efficient.

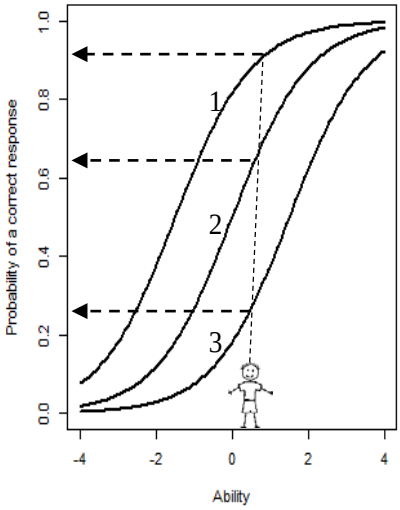
Item Characteristics: Difficulty, Guessing, and Discrimination

Most CATs use Item Response Theory (IRT) to dynamically adapt and individualize the assessment. This is done with iterative procedures to estimate student ability *after each response* and use that estimate to construct the best individualized test administration, which is a highly efficient approach to get a valid and reliable estimate of student ability. CATs can be 33% more efficient than non-adaptive measures.

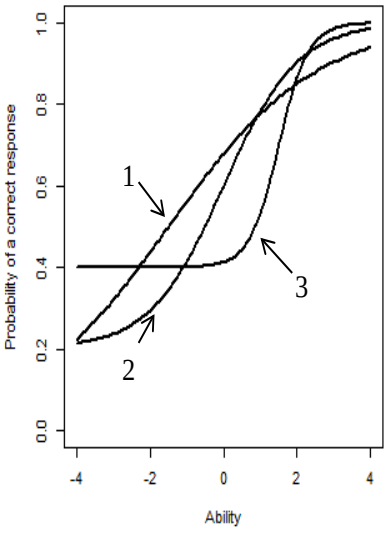
The properties of the items in the test bank are critically important; and there are three to consider.

1. **Difficulty**—defines the location of the item on the ability scale. Students with lower abilities are less likely to respond correctly to items with higher difficulty (range, -4 to +4).
2. **Guessing**—defines the probability that a low-ability student will guess correctly on an item (range, 0 to 99%). This is especially important to consider for multiple choice tests.
3. **Discrimination**—defines the amount of information an item provides at its location on the ability scale (range, 0 to 3). This is the discriminatory power of an item.

Some combination of these item characteristics define the *Item Characteristic Curve* (**ICC**) for every CAT item. **ICCs are used to make the test adaptive and generate student scores.**



The ICCs are used to estimate the probability of a correct response to each item during the assessment. Those probabilities are used to select the most appropriate and informative items. In the example on the right, the student has an ability in the middle of the scale, 0.5. The CAT will adapt to select the next item from among three choices in this example. The probably of a correct response on Item 1 (easiest item) is .92, Item 2 is .64, and Item 3 is .25 (hardest item). Item 2 is the best for this student because it is neither too hard nor too easy. Unfortunately, the ICCs to the right rely entirely on item difficulty and ignore the probability of a correct guess and the unique discrimination power of individual items.



In contrast, the image to the left displays the more complete type of ICC, which are used by FAST™. They incorporate the characteristics of difficulty, guessing, *and* discrimination. These are more precise and nuanced. Notice that unlike the ICCs above, the probability that a low ability student (4.0) will guess correctly on difficult items are well above zero. That is because CATs use multiple choice items. As illustrated above, most other CATs ignore guessing. Also notice the steepness (or slope) of the ICCs vary across items (left). That variation comes because the unique discrimination of each item is modeled. Most CATs ignore the unique discrimination of items. **CATs in FAST™ use more complete and precise psychometric models**, which is an important difference.