K-12 BLUEPRINT

A planning resource for personalizing learning

Computer Adaptive Testing

In 2010, two state consortia competed for the bulk of the $350 million dollars in assessment funding provided by the American Recovery and Reinvestment Act. Both planned to administer computer-based tests. However, the Smarter Balanced Assessment Consortium proposed the use of "computer adaptive" technology, while the Partnership for the Assessment of Readiness for College and Careers (PARCC) chose to use "fixed item" tests.

In a "fixed item" test, all students are asked to answer the same collection of questions, with varying levels of difficulty. Every test has the same number of items, and students must commit the same amount of time to the testing, regardless of how quickly they complete the test sections. In a "computer adaptive" test (CAT), students receive a tailored selection of questions, with the level of difficulty determined by their performance. The number of questions, length of time required to complete the testing, and overall difficulty of the test all differ from one student to another.

The CAT software estimates the ability level of the student based on his or her responses to each test item, refining the estimate over time. It then selects items based on the current ability estimate. There are typically four elements that work together to make this happen:

· Bank of test items - a large collection of field-tested questions spanning across all difficulty levels

· Test blueprint - a plan for how to select the questions so that the test contains the desired content and quantity of test items

· Algorithms - a series of steps that tells the computer how to select questions and when to end the test

· Algorithm constraints - a table of rules that inform the algorithm as it constructs the test

A student's score on an adaptive test is based on the difficulty of the items answered correctly or incorrectly, and not on the number of correct answers. This means that two students who answer the same number of questions correctly may receive different scores, because one student was presented with more challenging questions than the other.

Potential of CAT

· Shorter testing times reduce test-taker's fatigue

· No need for trained administrators

· Individually-paced, meaning test-takers don't need to wait for others to finish a test section before proceeding to the next one

· Accurate scores across a wide range of abilities, not only for average test takers

Limitations of CAT

· Not applicable to all subject areas and skills

· Require test-takers to be computer-literate

· Constrained by the number of, and access to, computers