

# RESEARCH BRIEFS

*Brief #17*  
Winter 2012

## Computer Adaptive Testing

With Computer Adaptive Testing (CAT) students sit before computers and take tests that are tailored to their ability levels. CAT works in the following way. If a student answers a particular question (item) correctly, the student then is given a more difficult item to answer. Conversely, if the student gives the wrong answer, an easier item will be presented. As one paper explains:

“This process continues until either the computer has enough information to produce a reliable test score or the test taker has reached the maximum number of items to be administered, whichever comes first. As simple as this may sound, the science behind CAT is anything but basic. Only through the use of a complex series of analyses, sophisticated algorithms, and large item pools is a test able to be administered in CAT format”<sup>1</sup>

### Item Response Theory

In order to customize a test to each student’s ability level, there must be a large collection (bank) of high quality items from which to choose. Item Response Theory (IRT) typically is used to create these banks. IRT is a mathematical procedure that “calculates” (determines) for each item the probability that a student with a given ability level will be able to answer it correctly. This makes it possible to customize a test that is targeted to the ability level of each test taker.

Calculating the characteristics of each item in the collection is an expensive task. Initially, every item must be administered to samples of students. Over time, items are added to the collection by having students who are being tested answer “untested” pilot items. Their answers to these items do not affect their test scores but their answers are used to calibrate the necessary item characteristics.<sup>2</sup>

# RESEARCH BRIEFS

## CAT Advantages and Disadvantages

There are a number of advantages of CAT, including quick turnaround of results and less time spent testing because the tests tend to consist of fewer items than found on most paper and pencil tests. Test security also is less of an issue because each student takes a different test. Of course, there will be some items that are asked of all students having the same ability level; however, the overall configuration of each test will differ for each student.

Most importantly, CAT provides a more accurate measure of a student's knowledge and skills than a paper and pencil test which typically includes many items that are too difficult for low ability students (who guess at many answers) and too easy for students of high ability, who are not challenged by many of the items.

As for disadvantages, there is the need for a large bank of calibrated items which can be used to construct tests. Students also need access to computers. Having an adequate number of computers for testing is a problem in many school districts. Regarding the actual test-taking, students cannot check previous answers because items must be answered in sequence. Students also cannot use a common test-taking strategy—skipping about the test and trying to answer the easier items first before moving on to the more difficult ones.

## MAP Tests

There are numerous organizations that offer CAT. Measures of Academic Progress (MAP) is one of the most popular and is used by many districts throughout Wisconsin and the United States. MAP was developed by the Northwest Evaluation Association.<sup>3</sup>

## Notes

<sup>1</sup> See Computer Adaptive Testing (CAT) in an Employment Context by Fetzer, Dainis, Lambert, and Meade: <http://www.previsor.com/pdf/WPCAT.pdf>, p. 1. Also see: <http://www.rasch.org/memo69.pdf>.

<sup>2</sup> For those who are interested, an excellent IRT interactive tutorial can be found online: <http://www.creative-wisdom.com/multimedia/IRTTHA.htm>.

<sup>3</sup> See <http://www.nwea.org/products-services/computer-based-adaptive-assessments>.