

Computer-adaptive testing: A new breed of assessment

BETH RUIZ, PhD; POLLY A. FITZ, MA, RD; CINDY LEWIS, MS, RD; CHRISTINE REIDY, RD

Once the province of scientific research and industrial manufacturing, computers are now more of a necessity than a luxury. Use of high-tech applications ranges from children's toys to the way we pay our bills, plan our vacations, and, yes, take tests. The proliferation of microchip technology has made computer-based testing a viable assessment option. More and more professionals (eg, architects, security dealers, nurses, and medical technicians) are seeking or implementing on-line options for testing. College students across the country can take computer-based tests for course placement, and prospective graduate students can opt to take a computerized admission examination rather than the traditional paper-and-pencil version.

EXPLORING COMPUTER-BASED TESTING

As early as the 1980s, the Commission on Dietetic Registration (CDR) began to explore the question of whether computer-based testing could be used for the registration examinations for dietitians and dietetic technicians. At that time, a changeover to the new technology was not pursued because of the inaccessibility of test centers with computer facilities. Since then, there has been considerable growth in and dispersion of viable testing centers that provide comfortable, secure, professional environments for computerized testing sessions. From the examinee's perspective, a computer-based examination offers many advantages: more frequent test administration, enhanced reliability of candidate performance, potential for reduced testing time, personalized test administration conditions, and more timely reporting of scores.

B. Ruiz is test development services manager, American College of Testing, Inc, Iowa City, Iowa; P. A. Fitz is a professor with the School of Allied Health Professions, University of Connecticut, Storrs, and chair of the Commission on Dietetic Registration (CDR); C. Lewis is the health promotion supervisor, St Peter's Community Hospital, Helena, Mont, and chair of CDR's Examination Panel; C. Reidy is with the Preprofessional Assessment Team of CDR, The American Dietetic Association, Chicago, Ill.

Address correspondence to: Christine Reidy, RD, Commission on Dietetic Registration, The American Dietetic Association, 216 W Jackson Blvd, Chicago, IL 60606-6995.

Because of the many advantages of computer-based testing, CDR has decided to implement computer-based testing for the entry-level dietitian and dietetic technician examinations by 1999. To facilitate understanding of the change from the currently used standardized testing format characterized by examination booklets, No. 2 pencils, and answer sheets to the 1999 format, which will feature individualized testing carrels and computer terminals, this article will discuss the most common types of computer-based testing and explain CDR's plans to administer computer-adaptive testing in 1999.

Fixed-Form Testing

The term *computer-based testing* is used to describe a broad spectrum of electronic assessments ranging from conventional multiple-choice tests administered on a personal computer to virtual reality simulations. For years, the most common type of computer-based testing has been fixed form, that is, a fixed number of items appearing in a fixed sequence. In other words, a conventional paper-and-pencil examination is transferred from the conventional test booklet format to a computer medium. One potential problem with fixed-form computer-based testing is reduced test security. Because the same test is offered every day, a candidate who takes the test on Monday could share information about the test with a candidate who will take it on another day.

Domain Sampling

An alternative method to meet this challenge in testing is to use domain sampling. In domain sampling, the computer accesses a large pool of items and chooses a predetermined number of items from each domain or content area. The computer is also instructed to target an average level of difficulty for the set of items that will make up each unique examination. Every candidate views the same number of items from each domain; however, each candidate's examination contains a different combination of questions. Although there may be some overlap, the goal is to minimize the number of questions any two candidates may share.

Computer-Adaptive Testing

A third type of computer-based testing is computer-adaptive testing (CAT). In contrast to computer-based fixed-form and domain-sampling tests, where the number of questions is preestablished, a computer-adaptive test is of variable length.

In CAT, some candidates may answer a minimum number of questions and others may answer many more. CAT takes advantage of recent developments in computer technology and item response theory to provide a customized, precise assessment for each candidate. An adaptive test is commonly shorter than the traditional paper-and-pencil, fixed-form, or domain-sampling examinations.

The two types of CAT — CAT for ability estimation and CAT for classification — are based on two distinct purposes in assessment. CAT for ability estimation spreads examinees across a continuum of ability, ranking test takers according to their relative levels of proficiency. This type of testing is common in college entrance examinations, scholarship award determinations, and general classroom evaluation for a grade. CAT for classification is used for assessments in which examinees are placed into one of two categories, such as pass (eg, ready to practice) and fail (eg, not ready to practice). The procedures, or algorithms, used by the computer to select items are very different in these two testing scenarios. The following examples from dietetics testing attempt to demonstrate the complex processes the computer uses to select items for each of these methods of CAT.

Computer-adaptive testing for ability estimation Suppose a training program decides to administer an end-of-semester test in clinical nutrition using CAT. Because the goal is to rank students according to their level of proficiency in the clinical nutrition course, CAT for ability estimation is selected. For ability testing, CAT usually begins with an item of moderate difficulty, such as this sample question: "The best way to assess dietary control of gestational diabetes is to monitor: (A) glycated hemoglobin, (B) urinary ketones, (C) serum glucose levels, (D) serum lipid levels." If the candidate selects "B," an incorrect response, an easier item will appear next. If the candidate selects "C," the correct response, the computer then chooses a more challenging item. The computer selects items on the basis of the previous response, continually evaluating the examinee's ability. Because computers are able to search through huge amounts of information quickly, the entire item pool can be searched to provide appropriate items. No fixed sequence of item presentation is followed because items are selected according to need rather than according to a preestablished order. When sufficient information has been obtained to identify the examinee's ability, the testing session ends.

Computer-adaptive testing for classification An example of CAT for classification would be an examination to determine whether a candidate for the RD credential is or is not ready to practice. Items available in the pool are selected according to their ability to discriminate between candidates in these two readiness categories. Through a field-test process, the items presented to the candidate have proven to be good at differentiating into which category an examinee is ultimately placed. Consider, for example, this rather difficult item: "Selection of a sensory evaluation instrument should be based on the: (A) research question, (B) experimental design, (C) amount of sample available, (D) number of panelists available." Candidates who fall into the ready-to-practice category choose the correct answer, "A," more often than candidates who are not ready to practice, but a right or wrong response does not drive the selection of the next question.

In CAT for classification, every candidate can potentially receive items of varying difficulty; however, each candidate's overall test is of similar difficulty to that of every other candidate. There are no difficult items for more able candidates or easier items for less able candidates. All candidates receive an

optimal set of items for classification purposes. The computer also selects items from all content areas of the domain of dietetics to ensure thorough content coverage. For instance, the current registration examinations call for a fixed percentage of the examination items to be in the area of nutrition services and a fixed percentage of the items to be in foodservice systems. In CAT for classification, candidates receive a proportionate number of items from each of the content areas that represent all aspects of dietetics as determined by a recent survey of practice. Each candidate receives a unique set of items sampling all content domains and can be confident that a representative sample of practice-based content has been presented in the items.

The advantages of computer-based testing include enhanced reliability of candidate performance and more timely reporting of scores

In computer-adaptive tests, the number of items presented during a testing session is determined by a set of stopping rules. These rules establish boundaries on the number of items presented, the length of a testing session, and the level of precision necessary for a final decision. For registration examinations, CDR will establish a minimum number of items a candidate must answer and the maximum duration of a testing session. For example, one candidate may only need to answer 100 items before the test is finished. Another candidate may need to answer 200 items before a pass or fail decision is made. Sufficient time will be provided to ensure adequate reading and response time for every candidate. Because CAT for classification offers solutions to problems inherent in fixed-form and domain-sampling testing and because it provides greater precision in measurement in a personalized environment, CDR chose this form of computer-based testing for its testing programs.

CDR's PLANS TO IMPLEMENT COMPUTER-ADAPTIVE TESTING

Over the next 4 years, CDR will be working with its testing agency, American College Testing, Inc, to carefully plan the implementation of CAT. Many questions must be answered during this time, including: How many months of the year will the examinations be administered? What will be the maximum and minimum test lengths? How soon will unsuccessful candidates be permitted to retest? How will candidates be informed of pass/fail status?

CDR's Examination Panel and its many RD and DTR test development committees, which represent diverse geographic locations and ethnic, racial, gender, and practice perspectives, will be working in cooperation with CDR's testing agency to answer these questions and to ensure that the validity and quality of the testing programs are maintained while enhancing service to the candidates. CDR is also mailing an informational diskette, "Computer Adaptive Testing: A New World of Options in Assessment of Dietetics Professionals," to dietetics education program directors and licensure boards. This interactive program provides further information on computer-based testing in a colorful and user-friendly environment.