NCSBN
National Council of State Boards of Nursing

111 E. Wacker Drive, Suite 2900
Chicago, IL 60601-4277
312.525.3600 NCSBN
866.293.9600 Toll-free
312.279.1036 NCLEX Examinations Department Fax

Computerized Adaptive Testing (CAT) Overview

Computerized adaptive testing (CAT) is a method for administering tests that merges existing computer technology with modern measurement theory to increase the efficiency of the testing process. The NCLEX® examination, administered via CAT, uses items with a variety of response formats, such as, single response multiple choice, multiple response, fill-in-the-blank, drag and drop, etc. and a variety of display formats, such as, chart/exhibit displays, tables, graphic images, etc. Regardless of the response or display formats, all items are scored as right or wrong; there is no partial credit.

With CAT, each candidate's test is unique. It is assembled interactively as the individual is tested. When the candidate answers an item, the computer recalculates the candidate's ability estimate based on all the responses including the most recent response. Next, the item bank which contains the available items classified by test plan content area and sorted by level of difficulty, is searched to find an item in the appropriate test plan area that best matches the candidate's ability. This item is selected and presented on the computer screen. This process is repeated each time an item is administered, creating an examination tailored to the individual's ability while fulfilling the NCLEX test plan requirements. The examination continues in this way until a pass or fail decision can be made. CAT provides greater measurement efficiency as it administers only those items which will offer the best measurement of the candidate's ability.

NCLEX examination decisions are not based on the number or percentage of items answered correctly, but rather on the difficulty of the items that a candidate can answer correctly 50% of the time. CAT administers test items with difficulty levels such that each candidate will answer about half correctly; these items provide the most information. Thus, all candidates answer about 50% correctly. Passing candidates answer 50% of more difficult items correctly, and failing candidates answer 50% of easier items correctly.

This is a fundamentally different approach to determining a candidate's ability as compared to traditional pencil-and-paper testing. Paper-and-pencil tests hold the difficulty of the examination constant (everyone receives the same items) and uses the variability in the number or percentage correct as the indicator that spreads out the candidates from low ability to high ability. CAT does the opposite. It holds the percentage correct near constant (50% on the NCLEX) and considers the difficulty of the items that can be answered correctly as the indicator that spreads out the candidates from low ability to high ability. CAT has the advantage of not wasting time by asking high ability candidates very easy items, which would provide very little information about their ability. Similarly CAT asks low ability candidates very few difficult questions, which would provide very little information about their ability, but would permit guessing to have a stronger influence.

Making CAT Work

NCSBN knows the precise difficulty of the approximately 1,700 to 2,000 items in each operational NCLEX examination item pool because each has been taken as a pretest item by hundreds of candidates and then

statistically analyzed. Picture the items all lined up, from easiest to hardest. If a candidate were asked the easiest items, he or she would get most of them right. If he or she were asked the hardest, they would probably get most wrong. As the candidate moves from easy to hard, there will come a point where the candidate goes from getting more right to getting more wrong. This is the point where the candidate is answering 50% correctly.

The candidate would probably answer items beyond this point (harder items) incorrectly. The candidate would likely get some right, but more wrong than right. Items before this point (easier items) would probably be answered correctly. The goal of CAT is to find that point for each candidate. That point is different for everyone. Nursing experts could probably answer at least half of the most difficult items correctly. Beginning nursing students would likely be able to correctly answer only half of the very easiest ones. Most candidates fall somewhere between those two extremes.

Once a bank of items with known item difficulties is established, a test can be administered. First, the computer asks a relatively easy item, and if the candidate answers it correctly, the computer selects a somewhat harder item. As the candidate continues answering correctly, the items get harder and harder. When the candidate starts missing questions, the items get easier until the candidate starts answering them correctly again, then the items begin to get a little harder. Each time the candidate answers one correctly, the next is harder. Each time the candidate answers one incorrectly, the next is easier. This zigzag process continues to narrow in on the point where the candidate answers 50% correctly, e.g., one right, then one wrong. That point represents the candidate's ability estimate. This is why everyone ends up correctly answering about 50% of the items he or she is asked.

After the candidate has answered the minimum number of items, the computer compares the candidate's estimated ability level to the passing standard and makes one of three decisions:

One, if the candidate is clearly above the passing standard, the examination ends, and the candidate passes.

Two, if the candidate is clearly below the passing standard, the examination ends, and the candidate fails.

Three, if the candidate's ability estimate is too close to the passing standard to determine with 95% certainty whether the candidate should pass or not, the computer continues to administer items.

Conceptually, "clearly" passing or failing is defined as when the confidence interval surrounding the candidate's ability level falls entirely above or below the passing standard. This confidence interval is the region within which the candidate's ability estimate could vary 95% of the time if the candidate answered more questions. This confidence interval shrinks a little after each question because the ability estimate is then based on more information. After each item, the candidate's ability level and the confidence interval are recomputed, including the most recent responses. When the confidence interval is entirely on one side or the other of the passing standard, the examination ends.

Of course, some people's ability level will be very close to the passing standard and all the items in the item pool might not be enough to make it "clear" whether they should pass or fail. When a candidate's ability level is very close to the passing standard, the computer continues to administer questions until the maximum number of items is reached. At this point, the computer disregards the confidence interval and considers only the final ability estimate. If it is above passing, the candidate passes. If it is equal to or below the passing standard, the candidate fails.

Even though candidates may answer different items and different numbers of items, the NCLEX examination administered using CAT is fair to every candidate. All examinations conform to the appropriate test plan (either the NCLEX-RN® or NCLEX-PN® Test Plan), which controls inclusion of important nursing content. All candidates have ample opportunity to demonstrate their ability, as the examination does not end until a stable pass/fail result is determined or time runs out.