Examining the Relationship between Student Scores on the National Council Licensing Examination for Registered Nurses (NCLEX-RN) and the Computer Adaptive Test (CAT)

by Barbara B. Laird

Return to Contents Page

Abstract
Laird studies the relationship between two computerized nursing tests and finds a relationship between the two sets of scores.

Nature of the Problem

In 1998, the nursing faculty at John Tyler Community College (JTCC) noticed a declining pass rate of our nursing graduates on the National Council Licensing Examination for Registered Nurses (NCLEX-RN). The pass rate for the JTCC nursing graduates had dropped to 86%, reflecting a downward spiral over the previous three years from a level of 100% (NCLEX-RN Profile Report, 1998).

As a result, we decided to require nursing students in their last nursing course to take the Arnett Development Corporation's Computer Adaptive Test (ADC-CAT). We chose the ADC-CAT because the company claimed that the test results would predict whether nursing students would pass or fail the NCLEX-RN. This assessment would allow faculty to identify students who were at risk for failing boards prior to graduation in order to assist these students with planning preparation for the NCLEX-RN, thereby also preventing the pass rate on the NCLEX-RN from declining further. The nursing faculty also wanted to provide the students with an opportunity to take a computerized adaptive test prior to taking the NCLEX-RN.

The ADC-CAT is an on-site computer-generated multiple-choice test that is based on the NCLEX-RN. The ADC-CAT models its practice test on the NCLEX-RN, which tests basic competencies that are essential in meeting the health care needs of patients. The examination is adaptive, individualizing the test to the student's ability, basing the number of questions a student answers on how the student answered the preceding question. Like the NCLEX-RN, the ADC-CAT is structured to allow a maximum of five hours to complete and a student may answer as few as 75

questions or a maximum of 265.

At John Tyler, the ADC-CAT is now mandatory for all senior nursing students, giving the students an opportunity prior to taking the national licensing examination to practice with a simulated computerized licensing examination that provides feedback on strengths and weaknesses regarding one's knowledge of nursing and critical thinking skills.

Hypothesis and Overview

Our research hypothesis was that there is a relationship between the scores on the ADC-CAT and the scores on the NCLEX-RN.

Until the 1990s, paper and pencil tests had predominated in nursing education. According to Anna (1998), testing on computers in nursing education is an emerging application. Computerized testing has been found to be more efficient in that students take less time to test (Olsen, 1990; Overton & Harms, 1997). Research has shown that anxiety based on computer adaptive testing decreases with hands-on practice on the computer (Fusqard, 1999).

Most nursing programs still administer paper and pencil tests to students despite the fact that the NCLEX-RN has been computerized since 1994. Anna (1998) states that the use of computerized testing in nursing education is relatively new. Anna implemented computerized testing in a nursing program with reported positive results from students who preferred the computerized method of testing. The study reported that students felt more prepared for the NCLEX-RN because of their exposure to computerized testing during their educational program. Prior to the study by Anna (1998), there was only one published study in the nursing literature that examined nursing students' performance on computerized tests. Bloom and Trice (1997) conducted a study comparing nursing students' performance on a paper and pencil test with those who took a computerized test. Their findings noted that the students who took the computerized test did just as well as those who took the paper and pencil test.

Halkitis and Leahy (1993) emphasized the importance of nursing students practicing with computerized testing prior to taking the licensing examination for nursing. They stated "being at ease with a computer will be half the battle" (Halkitis & Leahy, 1993, p. 384).

Also, the CAT test items are independent of each other. All information needed to answer the questions is

presented on a single screen. The examinees do not need to worry about remembering what a previous question asked or if there was something they needed to refer back to in that section.

Since the CAT is individualized, the length of time the examinee takes to complete the test varies. Generally, the CAT requires less time than traditional testing methods (Bloom & Trice, 1997; Overton & Harms, 1997; Forker & McDonald, 1996; Wise & Plake, 1990). How well the examinee performs on the test is not directly related to how fast the test is completed (Halkitis & Leahy, 1993).

In format, the ADC-CAT is identical to the NCLEX-RN and is set up and administered to students using lap-top computers. The test is adaptive in that it is individualized for each student. Questions are determined by how the student answered the previous question. "If the student gets the item correct, the next item is more difficult" (Beare, 1995, p. 35). The computer determines the total number of questions a student answered correctly and the number of questions answered incorrectly. The test measures students' performance against the entry level skills and job activity statements determined by the NLCEX-RN (Arnett Development Corporation Onsite Testing, 2000).

The test's three divisions are nursing process, categories of client needs, and cognitive levels. Any division where the student has missed 40% of the questions needs to be examined. The Arnett has eleven subcategories that test entry level skills based on a job analysis done by the National Council.

One division of the ADC-CAT covers the steps of the nursing process. The assessment component contains content that focuses on "nursing assessment, history, physical exam, documentation methods, and management of care" (Arnett Development Corporation School Report, 1998-2000, p. 4). The analysis component identifies needs, problems, and nursing diagnosis. The next component in the nursing process is planning. This section focuses on deciding what goals are a priority for a client and what interventions are necessary to meet the goals. Delegation and coordination skills are also a part of the planning component. The implementation component under the nursing process division focuses on carrying out the nursing interventions that were planned. The final component to this division is evaluation. This section focuses on the patient's progress and whether the goals are being met (Arnett Development Corporation School Report, 1998-2000, p. 4).

The second division of the ADC-CAT examines categories of client needs. There are four subcategories to this division: (1) safe, effective care environment, (2) physiological integrity, (3) psychosocial integrity, and (4) health

promotion and maintenance (Arnett Development Corporation School Report, 1998-2000).

The safe effective care environment includes physical and emotional safety and management of the care environment. The physiological integrity deals with basic medical-surgical nursing and the skills to provide the care. The psychosocial integrity includes psychiatric-mental health concepts and any threat to the psychosocial needs of the client.

The final category in this division is health promotion and maintenance. This area focuses on "helping clients attain their highest level of wellness and staying well" (Arnett Development Corporation School Report, 1998-2000, p. 6).

The ADC-CAT is also broken down by age, setting, and systems. The test questions are categorized by "standard age groups: newborn, infant—under one year, toddler—one to two years, preschool—three to five years, school age—six to eleven years, adolescent—twelve to eighteen, adult—nineteen to sixty-five" (Arnett Development Corporation, 1994, p. 2). According to the ADC, more questions focus on the adult and the older adult because that reflects "the current population and the health care environment" (Arnett Development Corporation, 1994, p. 2).

The test questions also simulate the "settings that nurses work in: hospital, nursing home, clinic, doctor's office/home health, and general" (Arnett Development Corporation, 1994, p. 2). Most questions are set to the acute care setting (hospital) because it is where the largest number of new graduates work.

The final division examines systems of the body: cardiovascular, endocrine, gastrointestinal, reproductive, integumentary, musculoskeletal, immune, neurosensory, psychological, renal, and respiratory.

The NCLEX-RN is a computerized licensure examination given nationally to graduates of nursing programs. "In 1992-1993, a job analysis study of newly licensed, entry-level registered nurses identified job activities in relation to frequency of performance and impact on maintaining client safety in various settings" (Beare, 1995, p. 33). In other words, essential competencies were identified that nurses had to be able to perform in their work settings (Norman, 1999; Beare, 1995; Bosma, 1994). The National Council of State Boards of Nursing developed a blueprint from the job analysis from which the licensure examination was developed. The job analysis is based on entry-level practice (Ridenour, 1999; NCLEX-RN Test Plan, 1999). The examination reflected the "knowledge, skills, and abilities essential for application of the nursing process to meet the needs of clients with commonly occurring health

problems" (Beare, 1995, p. 33).

The NCLEX-RN became computerized in 1994. Instead of a paper-and-pencil test that took two days, the NCLEX-RN became a computer adaptive test given in testing centers throughout the nation. The NCLEX-RN is offered six days a week and takes no more than five hours to complete with the average testing time frame being one hour and 45 minutes. The minimum number of questions a candidate must answer is 75 and the maximum is 265. The number of questions a candidate must answer is based on how the candidate answered preceding questions. The candidate cannot go back to previous questions and make changes in responses.

The NCLEX-RN is a multiple-choice examination. The first question given to a candidate is of low to moderate difficulty. If the candidate answers the first question correctly, then the computer will increase the difficulty of the test questions. However, if the candidate does not answer the first question correctly, then the computer will give the candidate an easier question. When the computer program has determined ability level and the candidate has answered a minimum of 75 questions, then the computer will automatically shut off. If the candidate has reached the five-hour limit, the computer will shut off and base whether the candidate has passed or failed on the answers of the last 60 questions (Beare, 1995).

The NCLEX-RN has four major categories of client needs which are further subdivided into ten other subcategories:

- Safe, effective care environment,
    - management of care
    - safety and infection control;
- Health promotion and maintenance
    - growth and development through the life span
    - prevention and early detection of disease;
- Psychosocial integrity
    - coping and adaptation
    - psychosocial adaptation;
- Physiological integrity
    - basic care and comfort

- o   pharmacological and parental therapies

- o   reduction of risk potential

- o   physiological adaptation (NCLEX-RN Test Plan, 1999).

Incorporated into the NCLEX-RN are nursing concepts and fundamental processes. These are integrated within the four major categories of client needs. The concepts and processes are nursing process, caring, communication and documentation, cultural awareness, self care, and teaching and learning.

Methods of Data Collection

The nursing program contracted with the Arnett Development Corporation (ADC) to administer the ADC-CAT to the nursing students enrolled in their last semester. Students were charged a fee of $19.95. ADC representatives designed a testing center at John Tyler Community College with lap-top computers installed with the testing software. The ADC representatives maintained test security by providing two company employees to proctor the examination. The students were monitored throughout the test and were only allowed to leave to use the restroom. Students were to report to the testing room at 7:45 A.M. with a valid driver's license as identification. Course faculty provided the company with a class roster. Course faculty also attended to verify student identity.

The test began promptly at 8 A.M., with a maximum amount of time given to complete the test of five hours. The minimum number of questions a student had to answer was 75 and the maximum number was 265.

As students completed the testing and exited the testing room, they were provided with documentation of their test results including a score indicating the probability of passing the NCLEX-RN. The documentation also identified the students' strengths and weaknesses. Recommendations for topics to study to improve areas of weakness prior to taking the licensing examination were also identified. The ADC representatives met with students individually and discussed the test results in detail at the completion of the CAT. Test-taking tips were shared for each question that the student answered incorrectly. ADC also provided students with study guides outlining individual weaknesses (Arnett Development Corporation, 2000).

The ADC representatives then met with course faculty and the program head of nursing to discuss the results of the test. Student scores were reviewed, and students were identified as being able to pass or fail the NCLEX-RN.

A written report was sent to the school also reflecting the same information.

A database was compiled identifying students who completed both the ADC-CAT and the NCLEX-RN between the spring semester 1998 to the spring semester 2000.   The database included identifying information such as student names, year of birth, and scores (pass/fail) on the ADC-CAT and the NCLEX-RN. The database was examined for completeness of information.

The sample examined in this study was nursing students who were enrolled in the last nursing course in the nursing program at John Tyler Community College between the spring semester 1998 and the spring semester 2000. There were 27 students in the spring semester 1998, 32 students in the fall semester 1998, 35 students in the spring semester 1999, 25 students in the fall semester 1999, and 18 students in the spring semester 2000. The sample is exhaustive of all graduating seniors in the JTCC nursing program.

The data was collected from the Arnett Development Corporation School Reports. The reports indicated which students received a pass/fail score on the ADC-CAT. The ADC defined the determining baseline for pass/fail score as 0.4. Any score with a 0.4 and above was given a pass and any score below and including 0.399 received a fail. The range of possible scores was 1.00 being the highest passing score and .00 as being the lowest possible failing score. The NCLEX-RN scores are reported to the nursing program only as pass/fail. The data was compiled in a chi-square template and the test value scores for chi-square were determined.

Data Analysis

The data on whether a student passed or failed the ADC-CAT and the NCLEX-RN were analyzed using the chi-square statistical test. The data was placed into a chi-square contingency table to determine if there was a relationship between the scores on the ADC-CAT and the scores on the NCLEX-RN. In addition, the analysis of data was used to test the null hypothesis. The database consisted of student names, year of birth, the percentile ability indicating pass/fail on the ADC-CAT, and the pass/fail scores on the NCLEX-RN.

There was one null hypothesis in this study. It was that there is no relationship between the scores on the ADC-CAT and the pass rate on the NCLEX-RN.

There was one alternative hypothesis in this study. It was that there will be a relationship between the scores

on the ADC-CAT and the pass rate on the NCLEX-RN.

A level of significance was set at 0.5 which is acceptable for an ex post facto research design where all variables cannot be controlled. The null hypothesis was to be rejected if there was less than a 95 percent level of confidence.

A chi-square test of association was used in this study. It is a one-tailed test. A chi-square table was used to determine the critical value at the .05 level of significance with one degree of freedom. The region of rejection was determined to be a calculated chi-square.

Assumptions and Limitations

The following assumptions were made to control the validity of the study. It was assumed that the ADC-CAT was developed and validated by experts and that it was a reliable and valid test. It was assumed that the data collected in this study was gathered from the population using school reports sent to the nursing program by the Arnett Development Corporation (ADC) and the NCLEX. It was also assumed that the students in the study all met the same criteria in that all students enrolled in the last semester of the nursing program had to have received a grade of C or better in the previous nursing course. Also, it was assumed that the chi-square was chosen because it is the statistical test used to determine a relationship or an association.

It was also assumed that the results of the study would be valid for the nursing program at John Tyler Community College. Finally, it was assumed that the population studied was representative of all current and future students enrolled in the nursing program as well as the students who were enrolled in the last semester in the nursing program from the spring semester 1998 to the spring semester 2000.

The data collected from John Tyler Community College nursing program may not be generalized to other programs due to different demographics. Also, due to the type of data, the concept of a normal distribution does not apply and since the data is not of an interval level, a parametric test for correlation was not appropriate either.

The external validity could not be determined due to the ADC being uncooperative in providing the reliability and validity data to the investigator. The corporation was contacted by the investigator via phone and fax requesting

the information, but they did not send it. The study was also limited because extraneous variables that could possibly affect the success in the ADC-CAT and the NCLEX-RN could not be eliminated.

Results

A total of 137 John Tyler Community College nursing students took the ADC-CAT during the time period between and including the spring semester 1998 through the spring semester 2000 (27 in the spring semester 1998; 32 in the fall semester 1998; 35 in the spring semester 1999; 25 in the fall semester 1999; and 18 in the spring semester 2000). The computer adaptive test that was administered to the nursing students in the last semester of the nursing program was required.

The pass/fail scores for the ADC-CAT and the NCLEX-RN for these students were examined using the chi-square test of association. Out of the 137 nursing students, 112 students passed and 25 students failed the NCLEX-RN, and 88 students passed and 49 students failed the ADC-CAT. The investigator examined how many students who passed the ADC-CAT also passed or failed the NCLEX-RN. Table 1 denotes this information. The computed chi-square was 4.6277 with one degree of freedom with a p-value of 0.0314.

Table 1

Comparison of the Number of Nursing Students Who Pass/Fail the NCLEX-RN and ADC-CAT

Pass ADC-CAT and Pass NCLEX-RN	71
Pass ADC-CAT and Fail NCLEX-RN	10
Fail ADC-CAT and Pass NCLEX-RN	41
Fail ADC-CAT and Fail NCLEX-RN	15

_____

Based on the results of the inferential statistical test of the chi-square test of association, the null hypothesis which states that there is no relationship between the pass/fail scores nursing students received on the ADC-CAT and

the pass/fail scores on the NCLEX-RN was rejected. The alternative hypothesis which states that there is a relationship between the pass/fail scores the nursing students received on the ADC-CAT and the pass/fail scores on the NCLEX-RN was accepted.

Table 2

Results of the Chi-square Used To Determine A Relationship Between the Scores On the ADC-CAT and the NCLEX-RN

Computed Chi-square	4.627773447
Degrees of freedom	1
p-value	0.031458367
Level of significance	0.05

Conclusions

The results of this study are unique in that there is not a published study examining nursing students' scores on the ADC-CAT and the NCLEX-RN to determine if there is a relationship. There were only two published studies in the nursing literature on computerized testing in nursing education. Anna (1998) and Bloom and Trice (1997) conducted studies on computerized testing in nursing. These studies indicated that students found computerized testing to be a positive experience and that student performance on the computerized test was the same or even better than paper and pencil tests.

This study examined the scores nursing students received on the ADC-CAT and the scores the students received on the NCLEX-RN. The results did show a relationship between the two sets of scores. Kirkpatrick, et al. (1996) states that "testing is a critical element in the educational process" (p. 113). Testing in the nursing program assures adequate practice test taking prior to graduation. The nursing board requires testing in order for a nursing graduate to become and practice as a registered nurse.

The results of this study indicated that there was a relationship between the scores received on the ADC-CAT

and the NCLEX-RN. Anna (1998) discussed that students felt more prepared for taking the NCLEX-RN because of their exposure to computerized testing during the educational process. This exposure to the ADC-CAT should have made students more comfortable in taking a computerized adaptive test; however, with the declining pass rate, this may not be the case. Other factors may need to be considered.

Since both the ADC-CAT and the NCLEX-RN were identical, then a puzzling question was why has the pass rate of the nursing students who graduated from JTCC continued to decline. Since there was a relationship between the two scores, then the ADC-CAT scores should have predicted whether the students would pass or fail the NCLEX-RN. The ADC-CAT did provide an opportunity for the nursing students to practice with a computerized test. The literature agrees with computerized testing practice by stating that testing does decrease the student's anxiety and helps build confidence as noted in the literature (Fusqard, 1999).

This study was designed to assist the faculty in evaluating the effectiveness of the ADC-CAT. Until this study, the nursing program at JTCC had not evaluated the use of the ADC-CAT as a preparatory step for the NCLEX-RN. The faculty had assumed that because the ADC-CAT was identical to the NCLEX-RN that there would not be a continued problem with the pass rates on the NCLEX-RN. According to Beare (1995), if the computerized test is to predict the success of the nursing student on the NCLEX-RN, then the computerized test must follow the same blueprint. As the pass rate continued to decline, the atmosphere of the nursing program faculty and students was one of concern and frustration.

The nursing faculty noted that the results of this study indicated that there was a relationship between the scores on the ADC-CAT and the NCLEX-RN despite the continued decline of the passing rate on the NCLEX-RN. The ADC-CAT does provide the nursing students at JTCC and throughout the United States with an opportunity to take a simulated computerized adaptive test prior to the NCLEX-RN. However, because of the continued decline in the scores on the NCLEX-RN despite the requirement of the ADC-CAT, there was a need for further study of this problem. However, despite the results showing that there was a relationship between the two scores, concern should be raised that JTCC's pass rate has continued to decline.

Recommendations

Based on the study results, conclusions made and implications drawn, the following recommendations were

proposed. First, it was recommended that the faculty continue to require the nursing students take the ADC-CAT in their last semester of the nursing program.

Second, it was also recommended that further study be done to identify factors that may account for the declining pass rate on the NCLEX-RN. A third recommendation was to continue to track statistics on students' performance based on test results to determine if there is a change. A final recommendation was to disseminate the information gleaned from this study to the program head and the entire nursing faculty.

REFERENCES

Anna, D. (1998, July/August). Computerized testing in a nursing curriculum: A case study. *Nurse Educator, 23*, 22-26.

Arnett Development Corporation. (1994). South Carolina: Arnett Development Corporation.

Arnett Development Corporation Onsite Testing. (2000). Available: http://www.arnettdevcorp.com/onsitetesting.html.

Arnett Development Corporation School Report. (1998-2000). South Carolina: Arnett Development Corporation.

Beare, P. (1995, May/June). NCLEX-RN update:Helping your students prepare. *Nurse Educator, 20,* 33-36.

Bloom, K. & Trice, L. (1997, March/April). The efficacy of individualized computerized testing in nursing education*. Computers in Nursing , 15,* 82-88.

Bosma, J. (1994, March). New approach to NCLEX. *Nursing Health Care, 15*, 115.

Forker, J. & McDonald, M. (1996, July/August). Perspectives on assessment: Methodologic trends in the health care professions computer adaptive and computer simulation testing. *Nurse Educator, 21*, 13-14.

Fuszard, B. (1999). *Innovative teaching strategies in nursing.* Maryland: Aspen Publishers, Inc.

Halkitis, P. & Leahy, J. (1993, September). Computer adaptive testing: The future is upon us. *Nursing & HealthCare, 14*, 378-385.

John Tyler Community College Catalog. (1998-2000). Chester: John Tyler Community College.

Kirkpatrick, J., Billings, D., Carlton, K., Cummings, R., Hanson, A., Malone, J., Miller, A., Robinson, L., & Zwirn, E. (1996, March/April). Computerized test development software: a comparative review*. Computers in Nursing, 14,* 113-125.

 LaDuca, A. (1994, June). Validation of professional licensure examinations. *Evaluation & the Health Professions, 17,* 178-198.

McMillian, J. & Schumacher, S. (1997). *Research in education: A conceptual model.* (4th ed.). New York: HarperCollins College Publishers.

NCLEX Program Reports. (1998). Chicago: National Council of State Boards of Nursing, Inc.

NCLEX-RN Test Plan. (1999). Chicago: National Council of State Boards of Nursing, Inc.

Norman, L. (1999). NCLEX examination development process: Job analysis to test items. *Issues, 20, 2*.

Olsen, J. (1990, April). Applying computerized adaptive testing in schools. *Measurement and Evaluation in Counseling and Development, 23*, 31-38.

Overton, R. & Harms, H. (1997, Spring). Adapting to adaptive testing. *Personnel Psychology, 50,* 171-226.

Ridenour, J. (1999). President's message. *Issues, 20*, 2.

Wise, S. & Plake, B. (1990, April). Computer based testing in higher education. *Measurement and Evaluation in Counseling and Development, 23,* 3-10.

Zaglaniczny, K. (1996, February). The transition of the national certification examination from paper and pencil to computer adaptive testing. *AANAJ, 64,* 9-14.

Barbara Laird is a professor of nursing and the coordinator of the nursing program at John Tyler Community College.

Return to Contents Page