USING COMPUTER-ADAPTIVE QUIZZING AS A TOOL FOR NCLEX-RN SUCCESS

by

JILLYN NOLAND PENCE

FELECIA WOOD, COMMITTEE CHAIR
ANGELA BENSON
ANDRE DENHAM
MARILYN HANDLEY
JULIA PHELAN

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Education in the Department of Educational Leadership in the Graduate School of The University of Alabama

TUSCALOOSA, ALABAMA

2016
ABSTRACT

A common goal among nurse educators is to adequately prepare graduates to be competent, caring professionals who deliver the highest standard of care. For the nursing school graduate, the first step to attaining this goal is success on the NCLEX-RN. Nurse educators are constantly seeking new strategies to help better prepare graduates for first time NCLEX-RN success. The purpose of this study was to examine the relationship between student access to computer-adaptive quizzing software program during the final semester of a baccalaureate nursing program and first time NCLEX-RN success.

Employing the principles of retrieval practice theory, computer-adaptive quizzing is a strategy used in nursing education to allow students to study, evaluate, remediate, and reevaluate mastery of nursing concepts. Computer-adaptive quizzing was developed using the foundational principles of computer-adaptive testing and item response theory, which have been utilized in education, psychology, and computer science for many years.

A retrospective, descriptive correlational design was used to compare NCLEX-RN outcomes of one cohort of students who did not have access to the computer-adaptive quizzing software (n=99) and one cohort who did have access (n=96). The computer-adaptive quizzing software program that was used in this study was PassPoint®, a product of Wolters Kluwer Health. There were no statistically significant differences between the two cohorts related to age, gender, race, pre-nursing science GPA, nursing GPA, nursing course failures, and if other degrees were held. There were no major course changes or revisions over the four semesters of the study other than the implementation of the PassPoint adaptive quizzing software. Chi-square
analysis suggested that there was an association between having access to and using the software and NCLEX-RN success (p< .001; df=1) with 16.16% of those not having access being unsuccessful on the NCLEX-RN compared to 1.05% of those with access being unsuccessful. Logistic regression showed the predictive model significantly predicted that 18% of the variability of NCLEX-RN success could be linked with usage of the adaptive quizzing software.
DEDICATION

This dissertation is dedicated to my family and friends who have endured the triumphs and the tears throughout this process. In particular, to my husband Ron, my biggest cheerleader, my rock, and best friend, and to my two boys, Will and Sam, who have patiently supported me and loved me throughout this stressful time in all of our lives. I love all of you very much and wouldn’t have made it without you.
LIST OF ABBREVIATIONS

ATI  Assessment Technologies Institute

BSN  Baccalaureate of Science in Nursing

CAQ  Computer-adaptive quizzing

CAT  Computer-adaptive testing

IRT  Item Response Theory

LSE  Linear standardized exam

NCLEX-RN  National Council Licensure Examination for Registered Nurses
I would like to thank the many colleagues, friends, and faculty members who have helped me with this research project. I would like to especially thank Dr. Felecia Wood for her guidance and expertise of the dissertation process. I would also like to thank my committee members Dr. Angela Benson, Dr. Andre Denham, Dr. Marilyn Handley, and Dr. Julia Phelan. Your suggestions, input, and guidance have helped me immensely and enriched the learning process. I’d also like to thank Dr. Erin Killingsworth and Dr. John Dantzler for their constant support and statistical expertise for this study. Many thanks to Dr. Cindy Berry for embracing the implementation of this study into her course as well as the important part she played in data collection. I’d like to thank Mrs. Cindy Ritter and Mrs. Jan Paine for helping to gather and record the extremely important demographical data. This research would not have been possible without each one of these people, and the support of my fellow graduate classmates, friends, family, and peers.
CONTENTS

ABSTRACT .......................................................................................................................... ii
DEDICATION ..................................................................................................................... iv
LIST OF ABBREVIATIONS ............................................................................................... v
ACKNOWLEDGEMENTS ................................................................................................... vi
LIST OF TABLES ............................................................................................................... x
CHAPTER 1. INTRODUCTION ............................................................................................. 1
  Statement of the Problem ............................................................................................... 2
  Purpose ............................................................................................................................. 4
  Theoretical framework ................................................................................................... 5
  Significance of the Study ................................................................................................. 8
  Definitions of Terms ....................................................................................................... 9
  Summary ........................................................................................................................ 10
CHAPTER 2. REVIEW OF LITERATURE .............................................................................. 11
  The NCLEX-RN ................................................................................................................ 11
  Predictor Variables ......................................................................................................... 13
  CAT .................................................................................................................................. 15
    Item - Response Theory ............................................................................................... 16
    Advantages .................................................................................................................... 17
    Disadvantages ............................................................................................................. 19
CHAPTER 3. RESEARCH METHODOLOGY ................................................................. 29

Research Design .......................................................................................... 29
  Setting ........................................................................................................ 29
Population and Sample .................................................................................. 30
  Procedures ................................................................................................ 30
Data Analysis ................................................................................................ 32
  Summary ................................................................................................... 33

CHAPTER 4. RESULTS ..................................................................................... 34

Research Questions and Hypotheses ............................................................ 34
Description of the Sample ........................................................................... 34
Data Analyses ............................................................................................... 40
  Usage ......................................................................................................... 40
  Access and Usage .................................................................................... 43
  Summary ................................................................................................... 44

CHAPTER 5. DISCUSSION, IMPLICATIONS, AND RECOMMENDATIONS ....... 46

Purpose Statement ....................................................................................... 46
Review of the Methodology ......................................................................... 47
Summary of the Results ............................................................................... 48
Discussion of the Results ................................................................................................................. 49
Limitations ....................................................................................................................................... 51
Implications for Nursing Education ................................................................................................. 52
Recommendations for Future Research ............................................................................................. 53
Study Summary .................................................................................................................................. 54
REFERENCES ....................................................................................................................................... 56
LIST OF TABLES

1. Chi-square Tests for Gender .................................................................36
2. Chi-square Tests for Race .................................................................37
3. Chi-square Tests for Nursing Course Failures ........................................37
4. Chi-square Tests for Other Degrees Held .............................................38
5. Levene’s Test for Equality of Variances for Age, Science GPA, Nursing GPA and t-tests for Equality of Means .................................................................39
6. Cohort 2 Practice Question Data: Science GPA, Nursing GPA .............41
7. Cohort 2 Usage of PassPoint: Age, Gender, and Race .........................42
8. Cohort 2 Usage of PassPoint: Science GPA, Nursing GPA, and Nursing Course Failures …42
9. Chi-square Tests for Effect of CAQ on NCLEX-RN Success ..................43
10. Logistic regression of Use of CAQ Related to NCLEX-RN Success .........44
CHAPTER 1

INTRODUCTION

Using Computer-Adaptive Quizzing as a Tool for NCLEX-RN Success

In 2014, 143,020 graduates of nursing education programs took the National Council Licensure Examination for Registered Nurses (NCLEX-RN) for the first time. Of those graduates, over 35,000 were not successful on their first attempt, resulting in an 83% national pass rate for 2014 (National Council of State Boards of Nursing [NCSBN], 2015). These numbers do not include the students not granted permission from their institution to sit for the examination due to not fulfilling graduation requirements. Pass rate percentages are calculated at a national, state, and institutional level in the United States. The NCSBN (2015) reports pass rates as the number of candidates taking the NCLEX-RN and the percent passing the exam. Some state boards of nursing set an expected minimum first-time pass rate percentage for schools of nursing as a benchmark for regulation and funding. Pass rates also play a vital role in the institution’s accreditation process (Pennington & Spurlock, 2010). Although an 83% national passing rate may be viewed positively, there are thousands of students who were unsuccessful who would likely disagree. Having a first-time pass rate goal of 100% success on the NCLEX-RN pushes nurse educators to explore strategies to improve their program graduates’ likelihood of success.

New graduates are facing higher acuity patients in the clinical environment; more community-based care requiring increased autonomy, and predicted nursing shortages for the large baby-boomer population of patients (Herrman & Johnson, 2009; Theisen, Janelle, &
Sandau, 2013). In order to meet these and other demands, schools of nursing, students, faculty, employers, and the general public all have a vested interest in the success of nursing program graduates on the NCLEX-RN. The accreditation, reputation, and funding of nursing education programs rely heavily on the NCLEX-RN pass rate (Santo, Frander, & Hawkins, 2013). Faculty may feel a sense of responsibility when pass rates drop, questioning their personal teaching approach and abilities (Frith, Sewell, & Clark, 2005; National League for Nursing [NLN], 2012; Shultz, 2010; Simon, McGinnis, & Krauss, 2013). New graduates experience financial, low self-esteem, and employment implications from failure on the NCLEX-RN while the public at large desires confidence in the competence of their caretakers (Yeom, 2013).

**Statement of the Problem**

Nurse educators face significant challenges in the preparation of students to become nurses in a highly demanding career. Nursing students have various demographic differences, but each student has the ultimate goal of passing the NCLEX-RN and beginning a career in nursing. With an increasing diversity in the nursing student population, nurse educators are challenged to effectively meet each student’s learning needs (Pennington & Spurlock, 2010; Rassool & Rawaf, 2007). The instructional design, or how the curricula are developed and delivered, is an important component of student success in learning critical nursing content.

Currently, many schools of nursing are using linear standardized exams (LSE) and multiple-choice tests with a preset bank of questions throughout the curriculum for evaluation, with some tests used as high-stakes exams for progression and graduation. Student scores on these high-stakes exams are one strategy for predicting if students will be successful on the
NCLEX-RN. However, the current products being used as high stakes predictors do not follow the computer-adaptive testing (CAT) format of the NCLEX-RN and conflicting evidence on their predictive value. With nursing curriculum integration of CAT and quizzing tools, the students may have a higher percentage of being successful on first time NCLEX-RN attempts.

The NCSBN (2015) changed the NCLEX-RN from a paper-and-pencil LSE to a variable length CAT in 1994. Different from an LSE, a CAT can assess an individual’s ability during the exam and progress the examinee through a process of identifying and examining the level of knowledge of particular subjects (Al-A’ali, 2007). Although computer experience is not a requirement, experience with CAT allows the nursing student to become familiar with the test format as well as identify and improve upon areas of identified weakness through remediation (Shapiro & Gephardt, 2012; Vrabel, 2004). There is a vast amount of literature studying predictors of success in nursing education programs and on the NCLEX-RN, yet very few conclusions. There is a need for nursing education programs and students to research strategies for helping a diverse body of students achieve NCLEX-RN success. (Koestler, 2015; NLN, 2012; Romeo, 2013; Simon et al., 2013; Taylor, Loftin, & Reyes, 2014). Research is inconclusive on the current practice of using LSE-formatted predictive testing products. CAT has been demonstrated to be more accurate, time efficient, and increases self-assessment among examinees as compared to standardized testing, and therefore could possibly be a solution to struggling students.

Based on the CAT platform, computer-adaptive quizzing (CAQ) is an emerging resource used in nursing education to help prepare students and graduates to sit for the NCLEX-RN. CAQ software programs allow students to utilize an adaptive question bank, at a self-set pace, with
customized remediation plans. As opposed to a static, fixed-form LSE, CAQ follows the CAT model and delivers questions based on the student’s ability level determined by answers to prior questions (Cox-Davenport & Phelan, 2015). CAQ gives students the opportunity to engage in retrieval practice learning which has the potential to enhance performance on the NCLEX-RN (Phelan, 2012).

Purpose

The purpose of this study was to examine the relationship between student access to computer-adaptive quizzing software program during the final semester of a baccalaureate nursing program and first time NCLEX-RN success. For students who had access to the CAQ tool, usage was defined as the number of practice questions answered, average mastery level of practice quizzes, and average mastery level of proctored quizzes. Students who had access to the CAQ tool and those who did not were analyzed to determine if access could be correlated with NCLEX-RN success.

Research question: Does the use of a computer-adaptive quizzing program as a tool for NCLEX-RN preparation during the final semester of a BSN curriculum predict first-time NCLEX-RN success?

Hypothesis 1: Students who utilized computer-adaptive quizzing as a practice tool in addition to the proctored classroom quizzes will more likely successfully pass the NCLEX-RN on the first attempt.

Hypothesis 2: Students achieving an average mastery level of five or greater on two proctored CAQs will more likely successfully pass the NCLEX-RN on the first attempt.
Hypothesis 3: Having access to and using a computer-adaptive quizzing software program will improve nursing students’ success rates on the first time NCLEX-RN attempt.

Theoretical Framework

Dating back to the 1600s, the ideas of testing, or recitation, to improve retention have been widely studied in psychology and education (Dobson, 2013; Leeming, 2002; Roediger & Karpicke, 2006a, 2006b; Storm, Bjork, & Storm, 2010). The testing effect phenomenon purports that the retrieval of information from memory during testing produces a better retention rate than studying and restudying the same material for an equal period of time. (Roediger & Butler, 2011). Based on this phenomenon, retrieval practice theory explores the positive effects that testing can have on learning. Getting the student to secure information in the long-term memory will allow for easier subsequent retrieval and better academic performance (Roediger & Karpicke, 2006b).

Human cognitive architecture includes both primary and secondary knowledge as well as short and long-term memory. Primary knowledge is knowledge that is biological, or that which is innately developed, such as speech and listening; secondary knowledge is the knowledge that is learned specifically through instruction, such as reading and writing (Geary, 2008). Secondary knowledge is content taught to students that requires transfer from short-term to long-term memory. The ability to retrieve information from long-term memory involves the use of mediators, which are words or phrases that come to the forefront of the mind with certain cues (Pyc & Rawson, 2010). Retrieval is defined as the “active, cue-driven process of reconstructing
knowledge” (Karpicke & Blunt, 2011, p. 772), and it is through practice use of the mediators (or cues) and information retrieval that knowledge begins to cement into the long-term memory.

The educator’s goal is for students to learn and understand concepts and be able to retrieve relevant information when needed. Unfortunately, evaluation, or testing, is often thought of as the finale of learning for both the student and the educator: if the student gets the right answer, then instruction can move forward. Oftentimes, however, students merely learn enough to get through an exam and then struggle with the concepts later. In science education, complex concepts requiring critical thinking and a large amount of recall build upon one another, making it imperative for students to retain information in the long term (Dobson, 2013; Karpicke & Blunt, 2011). Empirical evidence states that use of retrieval practice through repetitive testing improves long-term retention of concepts (Karpicke & Grimaldi, 2012; Karpicke & Roediger, 2008; Pyc & Rawson, 2010; Roediger & Butler, 2011; Roediger & Karpicke, 2006b).

Roediger & Karpicke (2006a) studied how students retain information based on the method of studying short passages of basic scientific data. The students were divided into three groups, each consisting of variations of short study sessions and testing events. Group one had four study sessions (SSSS), group two had three study sessions and a testing session (SSST), and group three had one study session and three testing sessions (STTT). Following the completion of the sessions, students were tested after five minutes, two days, and one week. At five minutes, recall results were very similar for all groups with group one having a slightly higher recall percentage (81% vs. 75%). Group 2 had the highest recall percentage at two days (64% vs. 54%), and group three had the highest recall at one week (56% vs. 42%) suggesting that the
students who were forced to access the mediators more often during testing, had a distinct advantage in long-term retention.

An additional study building from the retrieval practice theory, studied the effect of giving feedback during the test versus after completion of the test (Argawal, Karpicke, Kang, Roediger, & McDermott, 2008). The authors found that by deferring feedback until the end of the exam and allowing students to then re-study yielded higher long-term retention rates than giving immediate feedback during the exam. The process of giving students feedback along with multiple testing sessions allowed students’ learning to be continuous and improved retention (Butler, Karpicke, & Roediger, 2007). When using exams with concrete right or wrong answers such as multiple-choice, ordering, or true-false, feedback is essential to the effectiveness of the retrieval process (Roediger & Butler, 2011). Repeat testing allows the students to learn through the retrieval process, while feedback assures that the student is learning correct material.

As soon as nursing students begin classes in the nursing curriculum, new concepts are continually being introduced requiring an enormous amount of retention. For students to be successful in nursing school, and pass the NCLEX-RN, knowledge gained must be applied to both didactic and clinical situations to help make critical decisions regarding patient care. Repetition allows the memory to be continually stimulated to reinforce knowledge and practices into long-term memory. Rather than studying in a passive manor, simply re-reading or re-listening to material, retrieval practice requires the learner to actively recall and apply information (Dobson, 2013). Exercising retrieval practice in a study and testing environment through a process similar to what graduates will experience while taking the NCLEX-RN
following graduation improves retention of important concepts (Karpicke & Blunt, 2011; Roediger & Butler, 2011; Roediger & Karpicke, 2006a; Roediger & Karpicke, 2006b).

**Significance of the Study**

For graduates to be successful on the licensure exam, understanding of concepts and phenomena is essential due to the complexity of the NCLEX-RN questions. Through repeated exposure to practice questions and scenarios, students begin to understand and apply concepts rather than simply memorizing the material. Computer-adaptive quizzing programs allow the student to practice, remediate, and evaluate knowledge, understanding, and comprehension (Phelan, 2012). Literature searches on NCLEX-RN success and pass rates yielded many results regarding evaluation and remediation techniques, predictors of success and failure, and high-stakes testing. However, research on the use of CAT and CAQ as a strategy for NCLEX-RN is very limited. Following release of the *Fair Testing Guidelines* by the National League for Nursing, nurse educators were called upon to explore strategies outside high-stakes testing to improve NCLEX-RN success (NLN, 2012).

Following the last NCLEX-RN test plan change in 2013, the school of nursing where these data were collected had a decrease in NCLEX-RN passing percentage from 96% in spring of 2012, to 94% in spring of 2013, to 83% in spring of 2014. Although 83% is equal to the national average, the school of nursing was unsatisfied with this outcome. Thus, the program began to explore resources to better prepare graduates to be successful on the NCLEX-RN. This program does not currently utilize high-stakes testing as an evaluative measure or a benchmark for progression; however, it used formative and summative LSE at the time of this study.
Definitions of Terms

Definitions of terms critical to this research are presented. Conceptual definitions represent evidence-based literature; operational definitions describe the concept’s use in this research.

- **Computer-adaptive testing (CAT):** The process of delivering questions based on a mathematical relationship between examinees’ ability and the examinees’ responses with scoring based on rules implemented during development of the exam (Al-A’ali, 2007). For the purpose of this study, CAT is the logarithm platform for PassPoint software.

- **Computer-adaptive quizzing (CAQ):** The process of repeated practice through adaptive quizzes that determine and continuously update the student’s ability level while delivering calibrated items with known difficulty parameters in a learning environment rather than a testing environment (Cox-Davenport & Phelan, 2015). For the purpose of this study, CAQ refers to the students’ usage of the PassPoint quizzing software.

- **Linear standardized exam (LSE):** multiple-choice tests with a preset bank of questions (Al-A’ali, 2007). For the purpose of this study, the linear standardized exams were the type of exams the students completed throughout their nursing education curriculum.

- **Mastery Level:** a measure of the average difficulty level of the questions a student answers in each client need/nursing concept category. (Cox-Davenport & Phelan, 2015; Phelan, 2012; Phelan, 2013). For the purpose of this study, the mastery level achieved by the students served as an indicator for needed remediation or competency of conceptual ideas. The recommended minimum mastery level of PassPoint that is given by Wolters
Kluwer Health is three, however, the faculty responsible for the course in this study required a minimum mastery level of five on the overall exam scores with a preference of mastery level 8, with 5 being the mastery level for this study.

- Pass rate: the number of candidates taking the NCLEX-RN and percent passing (NCSBN, 2014). For the purpose of this study, the NCLEX-RN pass rate was used as an outcome.
- PassPoint: A mastery driven, computer-adaptive product used for learning, evaluation, and remediation for students during the final semester of the nursing education program (Wolters Kluwer Health, 2013). For the purpose of this study, PassPoint was the computer-adaptive testing software used in the final semester of the nursing curriculum to facilitate student preparation for the licensure exam.

Summary

This study examined if utilizing a computer-adaptive quizzing program for repeated retrieval practice during the final semester of a BSN program was related to success on the NCLEX-RN. The foundational framework used was the retrieval practice theory based on the testing-effect phenomenon, which states that retention of knowledge is increased with repeated retrieval of mediators, or cues, through repeated testing instead of passive studying of the same material over the same period of time. This study addressed the lack of research on CAQ as a tool for NCLEX-RN success in nursing education.
CHAPTER 2

REVIEW OF LITERATURE

Today, nurse educators are faced with challenges of diverse student bodies, emerging and ever-changing technological resources, web-based education, limited clinical opportunities for students, and the demand for a more active learning environment (Frith et al., 2005; Herman & Johnson, 2009; Rassool & Rawaf, 2007). Accreditation standards and boards of nursing require schools of nursing to maintain high academic standards as well as high NCLEX-RN pass rates (Koestler, 2015; NLN, 2012; Romeo, 2013, Santo et al., 2013; Simon et al., 2013; Taylor et al., 2014). Maintaining an above average first-time pass rate on the NCLEX-RN and the ways of achieving it remain a research priority among nurse educators in the United States (Frith et al., 2005; Shultz, 2010).

The focus of this literature review was to examine research related to CAT, NCLEX-RN-preparation strategies, predictors of NCLEX-RN success, and the use of CAQ as a tool in nursing education for NCLEX-RN success. Research of CAT and CAQ in nursing education is limited; therefore research from other academic areas utilizing CAT and CAQ will be discussed.

The NCLEX-RN

The NCLEX-RN examination is administered to graduates of diploma, associate, and baccalaureate nursing education programs. Successful completion of the exam is a requirement for licensure as a registered nurse in the United States. The exam measures the ability of graduates to meet a minimal level of competency expected of an entry level, novice registered
nurse (NCSBN, 2015). Faculty in nursing programs recognize the importance of educating students above the minimum competency level, however, first time pass rates on the NCLEX-RN remain the beacon for program quality (Giddens, 2009; McGahee, Gramling, & Reid, 2010; Phelan, 2012; Simon et al., 2013). Every three years, through job analysis of demands in the healthcare delivery system, educational high school readiness assessment through the American College Testing (ACT) service, and current standards evaluation, the NCSBN evaluates and updates the content of the NCLEX-RN (NCSBN, 2015). In 2014, the United States national percentage passage rate of first-time NCLEX-RN candidates was 83%. Following the latest updates made to the exam based on the job analysis, baccalaureate nursing programs experienced an 11% decrease in pass rates while associate degree nursing programs had a 19% decrease between 2013 and 2014 (NCSBN, 2015). Educators are actively seeking strategies to help students succeed on their first attempt of the exam.

The NCLEX-RN is a variable length CAT with a 95% confidence interval with a minimum of 75 questions up to a maximum of 265 questions, 15 of which are pilot questions. The exam begins with a few easy questions gradually increasing in difficulty until a question is missed. Once the examinee misses a question, the program delivers a less challenging question. As the examinee continues to correctly answer questions, the difficulty of questions increases until one of the stopping rules is met (NCSBN, 2015). The stopping rules of the NCLEX-RN are (a) the testing system is 95% confident that the examinee has or has not met the passing standard, (b) the maximum amount of time of 6 hours has expired, or (c) the maximum number of questions has been answered. If time runs out, the last 60 questions are scored, and if one answer falls below the passing standard, the examinee fails. If all 265 questions are answered within the
time frame, the testing system assesses the confidence interval at the time and if it falls below the passing standard, the examinee fails (Cook, O’Malley, & Roddey, 2005).

**Predictor Variables**

As nursing demands continue to increase in the United States, the mandate on nursing schools to produce competent graduates that are successful on the NCLEX-RN has increased (Frith et al., 2005 & Yeom, 2013). Researchers have studied several predictor variables for NCLEX-RN success with few conclusive results. Vandenhouten (2008) stated that high cumulative ACT score, age on admission to nursing school, grade point average (GPA), and grades in certain nursing courses did have a correlation to NCLEX-RN success but did not predict failure. Humphreys (2008) stated that age at NCLEX-RN examination was a predictive factor of NCLEX-RN success while gender, ethnicity, and marital status were not strong predictors of success. Yeom (2013) discussed the vast amount of research that has been conducted on end-of-curriculum, comprehensive, standardized testing measures (often as high-stakes exams) as predictors, yet highlighted that very little research has been conducted on content-specific exams in order to identify high-risk students and remediate as necessary. In a study examining if content-specific standardized exam scores predicted NCLEX-RN success, Yeom (2013) stated that adult medical-surgical, pharmacology, and community health nursing were strong predictors of NCLEX-RN success while pediatric, maternal-newborn, leadership, and foundation scores did not have a strong predictive value.

In a study of two cohorts examining the ability to predict probability of NCLEX-RN success using standardized exams, Daley, Kirkpatrick, Frazier, Chung, and Moser (2003), reported there were two program variables associated with NCLEX-RN success: final course
grade of the senior didactic medical-surgical class and cumulative program GPA. The two exams used for predictability were the Health Education Systems Incorporated (HESI) Exit Examination and the Mosby AssessTest. Results of these standardized exams differed greatly between successful and unsuccessful students, with unsuccessful students having lower course grades and program GPAs. Demographic variables for the students who were successful included a higher average age (22.9 versus 20.4) at time of testing, higher average preprogram GPA (3.3 versus 3.1), higher ACT score (23.6 versus 19.6), and white students compared to all non-white students (33% versus 4%). Daley et al. (2003) recommended further research to substantiate demographic and program variables that could indicate high-risk students and allow for more proactive interventions to be done prior to the student taking the NCLEX-RN.

Beeman and Waterhouse (2001) studied predictors of NCLEX-RN success or failure with 289 nursing students over a three-year period. Results showed that students with lower nursing theory grades (C+ or lower) had a higher risk of being unsuccessful on the NCLEX-RN while students with higher grades in the core courses were more likely to be successful on the exam. In a study of 505 nursing students, Beeson and Kissling (2001) found that students with lower biology scores and those with lower grades during the sophomore year were more likely to be unsuccessful on the NCLEX-RN. Another significant predictor in this study was that students who earned one grade of C or below in a nursing course had a NCLEX-RN passing rate of 84% compared to that of students with As and Bs who had a 97% passing rate. The passing rate for students with three or more Cs or below in nursing courses dropped to 51% (Beeson & Kissling, 2001).
McGahee et al. (2010) studied predictor variables with results indicating that higher predictability of NCLEX-RN success or failure is found among interactive variables rather than any one isolated variable. Grades in the first semester of the nursing curriculum (health assessment, foundations, and pathophysiology), results of the RN Assessment Test, a standardized exam given in the final senior semester, and science GPA were highly correlated to the outcome of the NCLEX-RN. Following data analysis, it was found that the predictor variables were much more accurate in prediction of success rather than prediction of failure, mainly because the sample of failures was so much smaller than those who were successful.

Factors such as gender, race, age, and nursing GPA in relation to NCLEX-RN success were studied within a sample of 368 baccalaureate students over a ten-year period. Haas, Nugent, and Rule (2003) found that male students and African American students had a higher risk of being unsuccessful on the NCLEX-RN. Younger students had a higher passing rate while those who were predicted to be successful, yet were not, did not mirror these variables. Haas et al. (2003) suggests that students with any of the at-risk variables be included in remediation and offered proactive interventions to help prevent first-time testing failure.

**Computer-Adaptive Testing**

Computer-adaptive testing (CAT) is the process of delivering questions based on a mathematical relationship between examinees’ ability and the examinees’ responses, and scoring based on rules implemented during development of the exam (Al-A’ali, 2007). The major difference between a LSE and a CAT is that the CAT is individualized to the person taking the exam whereas the LSE delivers one set of questions to every student. By individualizing
questions to the person’s ability, competency of material is more easily assessed; having significance in nursing, since the level of difficulty of questions can increase to measure higher-level questions and problem solving skills (Hsu, Wang, & Chen, 2013; Wendt & Kenny, 2009). CAT is designed with a common starting point, a passing standard, stopping rules, and a confidence interval, all of which are decided upon during construction of the examination program (Cook et al., 2005; Hsu et al., 2013).

Once the student has met the set criteria to stop the exam, a score, commonly referred to as a mastery level is released to the student. The mastery level is determined by the difficulty level of the question, determined during development, and then calculated based on a student’s mastery of the presented information (Phelan, 2012).

**Item Response Theory**

CAT development is based on the Item Response Theory (IRT). Item response theory is a psychometric model based on the idea that an individual’s response is representative of the person’s knowledge and the item parameters (Al-Ali, 2007). The response of the individual determines how well one can pick the correct answer based on knowledge level (Gouli, Kornilakis, Papanikolaou, & Grigoriado, 2001; Kustiyaningsih & Cahyani, 2013). Through the use of IRT models, adaptive testing programs are able to delineate what questions will most accurately assess mastery of the examinee. A common IRT model is the Rasch model (or 1-PL model), which uses logits as the unit of measure to compare data and predicted responses (Lavin & Rosario-Sim, 2013; Yorke, Horton, & Jones, 2012). The NCLEX-RN is based on a Rasch unidimensional IRT model that looks at the single construct of nursing knowledge and
focuses on a single parameter of item difficulty (Cook et al., 2005; Kustiyahningsih & Cahyani, 2013; O’Neill & Reynolds, 2006).

Al-Ali (2007) used IRT to study measurement of students’ academic abilities in a mathematics course. Students took a conventional written exam where the final score received an average level of competency. For the students who proceeded to the adaptive testing method, questions began on the level determined by the conventional exam. Results showed that the computer-adaptive model required fewer questions to measure the students’ highest level of academic ability while not compromising the validity of the examination.

**Advantages**

There are many advantages to utilizing a computer-adaptive format. The primary advantage is the ability of the program to adapt to the examinee’s ability, providing a more accurate measurement of the competence of the examinee in fewer items and within a shorter timeframe (Challis, 2005). Rather than every student taking every question, CAT allows the student to only take the minimum number of questions required to achieve the set passing standard (Hsu et al., 2013; NCSBN, 2015). CAT is less tedious for the student because the test avoids questions that are too easy or too difficult for the student, therefore, not wasting time and requiring up to one-fourth of the testing time compared to paper-and-pencil exams (Gouli, Papanikolaou, & Grigoriadou, 2002; Vrabel, 2004). If fewer items are given to the examinee (decreased exposure), an added advantage is lowered risk of the test questions being revealed to the general population than if every student sees every question; test security is increased (Al-A’ali, 2007). Bulut and Kan (2012) compared CAT to the traditional paper-and-pencil admission test for students in Turkey applying for entrance into graduate school. Results indicated that
CAT gave 93% or higher accuracy measures with up to 70% fewer questions than the traditional testing method.

Kustiyahningsih and Cahyani (2013) assessed if the computer-adaptive format measured students’ academic abilities, and also compared pretest and posttest scores of students who took exams in a conventional method to those who used an adaptive method. In a study of two groups, each with 88 students, each student was given a pretest to determine existing knowledge level. Group one took a series of three exams using the CAT method, with the beginning content level being what was scored on the pretest. Group two took the same three exams, but in the conventional method with linear standardized questions. Results of the three adaptive exams showed test averages of 52.8, 76.4, and 86.1, respectively with conventional exam average scores being 55.3, 62.8, and 72.2. Scores on the pretest and posttests also differed between the two groups. The adaptive testing group’s pretest score average was 72.46 with the posttest score being 78.57 while the conventional method group’s averages were 72.67 and 76.65, respectively. From this study, conclusions include that using a CAT format gives the learner a more accurate challenge for their ability as well as provides additional learning opportunities as the student uses the adaptive program.

A second advantage of CAT is immediate and accurate feedback. Being able to have feedback immediately allows students to identify problem areas quickly. Shapiro and Gephardt (2012) reported that the CAT format identifies competent areas as well as deficient areas allowing students to remediate needed material. Although nurse educators play a major role in preparing students to be successful on both the NCLEX-RN and the nursing field, students need to be able to self-assess, identify areas of needed improvement, and make changes in their
personal learning. The LSE-format exams currently being used have this ability also, however if the student has a highly correct guess rate, deficient areas may not be accurately identified (Lau, Lau, Hong, & Usop, 2011).

A final advantage of CAT is flexibility and accessibility. With student diversity, non-traditional student enrollment, and web-based learning continuing to increase in nursing education, flexibility and accessibility are an ever-present factor in program selection. By using products with the computer-adaptive format, students can identify areas of deficiency, practicing and remediating at an individualized pace prior to taking the NCLEX-RN (Challis, 2005). In many situations, students are able to test in remote locations decreasing distractions, expenses, and time.

Disadvantages

There are also some disadvantages to CAT. The most significant disadvantage is cost. The start-up costs are extremely high: hardware, software, and programming. The larger the item bank is, the better the measurement of the examinee’s ability. A large item bank allows the computer to be more detailed in item delivery, yet it takes a significant amount of time and manpower to create one (Cook et al., 2005). Software development, or programming, is product specific. Therefore, the task of covering every idea available within a construct is massive, which drives up the cost. The delivery devices (computers, laptops, tablets) outside of those used for development training are usually bought by the consumer, but can also be cost-prohibitive depending on the socioeconomic level of the consumer. Matthews (2009) discussed the cost and importance of implementing security measures that will help to prevent outsiders from hacking into the testing programs. All of these costs will be passed onto the consumer, presenting a
challenge for exam developers to stay within reasonable expense parameters. Computer-adaptive software is programmed to build off each answer to the previous question; therefore a premise of the CAT products is that allowing examinees to revisit and change responses would jeopardize the validity of the exam results (Han, 2013). In a qualitative study looking at preference of paper-and-pencil testing, linear computer testing, and CAT, students discussed that another disadvantage was having the inability to skip or revisit questions (Vrabel, 2004). Papanastasiou (2002) reported that allowing examinees to have the option to return to previous questions could increase test performance by decreasing test anxiety that comes with knowing each answer is final.

**NCLEX-RN Preparation**

One of the main roles of nursing education programs is to prepare students for nursing careers, with the stakeholders holding the nursing programs accountable for student success (Frith et al., 2005; Herrman & Johnson, 2009; March & Ambrose, 2010; Yeom, 2013). In response to this demand, nursing programs have been urged to implement various strategies to improve success on the NCLEX-RN and to help predict the likelihood of failure, including implementation of progression and graduation policies utilizing high-stakes testing products to determine if students will graduate or take the NCLEX-RN (Santo et al., 2013, Shultz, 2010). Other strategies include increasing admission requirements, recommending or requiring review programs, and providing remediation and tutoring to students at risk of failing (Crow, Handley, Morrison, & Shelton, 2004; Herrman & Johnson, 2009).
High-Stakes Testing

High-stakes testing is defined in the *Standards for Educational and Psychological Testing* (1999) as “a test used to provide results that have important, direct consequences for examinees, programs, or institutions involved in the testing” (p. 176). Heubert and Hauser (1999) reported that high-stakes testing involves decisions regarding a student’s tracking, promotion, or graduation. Since nursing programs, program graduates, and institutions are directly impacted by outcomes of the NCLEX-RN, it can be classified as a high-stakes test. Nursing schools throughout the country have begun implementing high-stakes testing policies for progression and graduation that either prevent students from graduating or withhold permission to take the NCLEX-RN based on high-stakes exams given as part of the nursing program (Spurlock & Hanks, 2004). As part of these policies, faculty members predetermine an acceptable benchmark score on a standardized test product accompanied by policies regarding progression and graduation from the nursing education program. Use of these standardized exams as determinants of progression does not consider other factors contributing to the success of students. Many schools of nursing are adopting progression policies largely as a result of successful marketing of the companies that are making the exams rather than looking at all of the available research (Spurlock, 2006).

Following completion of the curriculum, diplomas are sometimes withheld based on the results of the high-stakes exams. Giddens (2009) believed these policies to be unfair and potentially an unethical educational practice. The National League for Nursing (2012) expressed significant concern that scores on standardized tests in nursing education are being used to block graduation or deny students eligibility to take the NCLEX-RN. Heubert and Hauser (1999) stated
that high-stakes tests can often be used in inappropriate ways, causing the focus of the education to be blurred from a quality program to one that focuses on exam success. Shultz (2010) argued that programs are rushing to implement progression policies based on products that they have not researched, possibly making decisions that don’t match the program philosophy. The author suggests that schools of nursing should possibly reexamine their curricula and ultimately decide if turning to commercial products “rather than our own rigor and expertise, to ultimately determine the student’s fate for progression and graduation” (p. 205) is the best way to evaluate students. Although the views expressed by the authors above frown upon high-stakes testing, the reality is that schools of nursing want graduates to be successful and will continue to implement progression policies based on high-stakes testing results.

Two predominant testing products used by nursing programs as predictors for NCLEX-RN success in the United States, the Assessment Technologies Institute (ATI) Comprehensive Exam package and the Health Education Systems Incorporated (HESI) Exit Examination (HESI E2), will be discussed.

The ATI comprehensive assessment and remediation package offers tutorials, proctored and non-proctored exams, critical thinking entrance and exit exams, a learning style inventory, content-specific exams, a comprehensive predictor exam, immediate feedback, and a remediation program. The exams are in LSE format and each student takes the same exams (ATI, 2014). Benchmarks based on company research are suggested by the company but are occasionally adjusted by the individual programs. The ATI exam products are used for feedback to the students and the faculty as well as identifying at-risk students, and are occasionally used as a requirement for progression and graduation (Davenport, 2007; Jacobs & Koehn, 2006).
The HESI testing program includes content specific exams and a cumulative exit exam. Benchmarks are set by the company and recommended to the nursing programs that ultimately make the final decision on how to use the results. The HESI E² is a 160-item exam that is in LSE format using the HESI predictability model to predict success on the NCLEX-RN. It is primarily used as a progression-to-graduation tool (Elsevier, 2014; Spurlock & Hanks, 2004).

The validity of the HESI E²’s ability to predict NCLEX-RN success has been studied in depth in many studies (Harding, 2010; Nibert, Young, & Adamson, 2002; Morrison, Adamson, Nibert, & Hsia, 2004). These studies discuss that the HESI is between 92-98% accurate in predicting NCLEX-RN success. While this looks very impressive, the argument that this is a skewed prediction of success rather than failure plays an important role in determining best practices in NCLEX-RN success preparation (Harding, 2010; Phelan, 2012; Spurlock & Hanks, 2004). If the research that is being done is only measuring the students that were successful on the high-stakes test, then the data fails to represent those students who may not have been allowed to take the NCLEX-RN. Nursing programs need to be focusing on predicting those students who are at risk of failing rather than using a tool to predict that the higher performing students are going to be successful. While the ongoing debate over the accuracy of these predictor exams continues, the importance of NCLEX-RN success remains a priority. CAT programs are an additional type of resource available for NCLEX-RN preparation and closely follow the format of the NCLEX-RN.

The high-stakes testing products predominantly in the literature are computerized, multiple-choice LSEs given in the last semester of the student’s nursing curriculum. The same question set is given to each student, regardless of ability, knowledge, or responses (ATI, 2014;
Phelan, 2012). Although these exams are commonly given on a computer, they are not individualized; the computer is merely an alternate format of delivery from paper and pencil. The majority of LSE questions used in nursing programs are multiple-choice, which follows a similar format of the NCLEX-RN, but are a lower difficulty level (Frith et al., 2005). Guessing is another factor of multiple-choice exams that poses a challenge to adequately assess true ability of the student. With a LSE, a student may guess on several items, skewing the results either positively or negatively, without the instructor’s knowledge leading to a false sense of security for the faculty and the student (Lau et al., 2011).

**Remediation**

A second preparation strategy reviewed in the literature is content remediation. Students may arrive at college deficient in basic skills and knowledge for being college success (Wiles, 2015). Bautsch (2011) states that a majority of high school graduates are in need of remediation of one or more of the main concepts of K-12 education: math, science, reading, and English. If students do not have mastery of the basic educational concepts, then the task for nursing education becomes greater. At-risk students must be identified early and given additional support to improve the chance of being successful (Corrigan-Magaldi, Colalillo, & Molloy, 2014). Wiles (2015) discussed that students must be aware of what they do not know in order to become actively engaged in the learning process. Rawson, Dunlosky, and Sciartelli (2013) studied two separate classes of introductory psychology students to examine if the practices of successive relearning using learning sessions and practice tests increased course outcome learning and long-term retention. Through this process, needed areas of remediation are identified and studied in order to understand the concept when seen on a later practice test or final evaluation. Results of
both studies revealed that spacing study sessions that use practice tests requiring long-term memory access through retrieval practice increased final evaluation performance and long-term retention.

Nursing students have common deficiencies in test taking skills and adequate resources for success and need to be given guidance and remediation early in the nursing curriculum (Wiles, 2015). Through faculty intervention and communication, students can learn effective study skills and test-taking strategies to improve performance during nursing school as well as on the NCLEX-RN (Corrigan-Magaldi et al., 2014).

**Computer-Adaptive Quizzing**

Computer-adaptive quizzing is an instructional technique being used more frequently in nursing education. The process and development of CAQ is based on the same foundation of the item response theory in CAT. Phelan (2012) reported that using a CAQ software program to develop quizzes “brings together best practices in learning, technology and assessment to provide a unique adaptive quizzing tool for students and instructors” (p. 1). The quizzing mimics testing but can be used for formative assessment by the instructor as well as learning and remediation by the student. Open access for practice gives students the option to utilize the CAQ software as often as needed and incorporate the feedback given through the software program into continued learning. As students use the adaptive program more and answer questions with increasing difficulty, the student can review both correct and incorrect answers with rationale (Wolters Kluwer Health, 2013). In a study of students’ usage of a CAQ program, increased usage was associated with increased mastery of content (Cox-Davenport & Phelan, 2015).
Through increased usage of CAQ, students have increased exposure to content as well as increased retrieval practice of important content. Corrigan-Magaldi et al. (2014) studied the effect of using an adaptive quizzing system with 11 at-risk nursing students. Students were required to complete weekly assignments in a NCLEX-RN CAT software program as well as utilize course resources. Faculty monitored activity and progress and provided encouragement and support through frequent communication. Results demonstrated that nine of the eleven participating students passed the NCLEX-RN, suggesting that the strategies used are effective with at-risk students.

Cox-Davenport and Phelan (2015) researched usage of an adaptive quizzing program in a school of nursing that had previously utilized a commercial standardized exam. Scores on the commercial standardized exam were previously used as guidelines to determine if students would be permitted to sit for the national board exam. The school of nursing sought to promote increased success on the NCLEX-RN, determine if the commercial exam was accurate in predicting NCLEX-RN success, and shift from a punitive outcome to a positive learning experience. The adaptive quizzing program was implemented in a NCLEX-RN preparatory class taken in the last semester of a student’s senior year across three cohorts. Data from the adaptive quizzing program usage, mastery level average, predictor exam results, and NCLEX-RN results were collected. Results showed that overall average mastery level was highly correlated to usage, yet due to a small number of failures, no analysis was conducted to determine mastery level correlation to NCLEX-RN success.

PassPoint is a formative, adaptive, quizzing program based on a mastery level algorithm, and contains over 10,000 questions covering topics aligned to the NCLEX-RN blueprint. It is a
resource to be used as a tool to help prepare students to take the NCLEX-RN (Wolters Kluwer Health, 2013). PassPoint was released to schools of nursing in fall 2013, with pilot data showing an increase in NCLEX-RN pass rates of students utilizing it (Phelan, 2012). The product can be used as a teaching strategy using both comprehensive and content-specific quizzes, to help identify areas of knowledge deficiency, as well as used as a remediation tool for at-risk students (Phelan, 2012; Wolters Kluwer Health, 2013). The question delivery algorithm presents students with quizzes composed of questions with a range of difficulty levels, establishing a mastery level at the conclusion of each quiz taken. The mastery level scale is 1-8, with one being simple and eight being very difficult. Phelan (2012, 2013) discussed that mastery level achieved on PrepU, the precursor CAQ product to PassPoint, has a direct correlation to students’ usage of the product and to NCLEX-RN success.

**Recommendations for Future Research**

Past research has studied predictor variables such as science GPA, ACT and Scholastic Aptitude Test (SAT) scores, reading scores, age, environmental stressors, high school education, and GPA in nursing courses, with very few of them showing a positive correlation with success on the NCLEX-RN (Frith et al., 2005; Simon et al., 2013). Research of utilization of commercial high-stakes exams shows conflicting results on the predictive ability of NCLEX-RN success. There is limited research available on best strategies for NCLEX-RN preparation; therefore, more studies are needed to determine strategies for increasing student success, identification of at-risk students, and the most effective preparation strategies for diverse populations (Corrigan-Magaldi et al., 2014; Davenport, 2007; Jacobs & Koehn, 2006; Lavin & Rosario-Sim, 2013; Romeo, 2013). Additional research on the ability to predict NCLEX-RN
failure rather than success is needed, but with numbers of failure being small, based on national averages, this remains extremely challenging (Crow, Handley, Morrison, & Shelton, 2004; Yeom, 2013). There is little research on the use of CAQ as a tool for NCLEX-RN preparation, therefore further research is needed to examine if utilizing a computer-adaptive model is an effective strategy for NCLEX-RN success (Phelan, 2012).
CHAPTER 3
RESEARCH METHODOLOGY

The purpose of this study was to examine the relationship between student access to computer-adaptive quizzing software program during the final semester of a baccalaureate nursing program and first time NCLEX-RN success. This chapter outlines the study process used by the investigator to determine if having access to a CAQ resource was associated with NCLEX-RN success. This chapter includes the research design, setting, population, sample size, procedures, data collection and analysis, and the chapter summary.

Research Design

This research study employed a retrospective, descriptive, correlational study design. This design allowed the investigator to examine relationships between students’ access and usage of a CAT program and first-time NCLEX-RN success.

Setting

The setting for this research study was a baccalaureate nursing program in a private Christian university located in a metropolitan city in the southeast United States. The university’s total enrollment was approximately 5200 students, with approximately 400 undergraduate nursing students. The school of nursing is part of the College of Health Sciences and is accredited by the Commission on Collegiate Nursing Education. Degrees offered in the school of nursing include baccalaureate degrees, master’s degrees, and doctoral degrees. Qualifications of the investigator included a Master’s of Science Degree in nursing education, 11 years of teaching
experience in nursing education, and 20 years of nursing experience. For this study, the investigator worked with the course coordinator of the critical care course in the final semester of the baccalaureate curriculum. The investigator did not have a teaching or evaluative role in the critical care course.

**Population and Sample**

The population for this study was comprised of last-semester baccalaureate nursing students in a critical care course from four academically sequential semesters. A purposive sample of de-identified nursing graduates’ records from the fall and spring semesters during the academic years 2013-2014 and 2014-2015 were used to examine if access to and usage of a computer-adaptive software program, PassPoint, was associated with first time NCLEX-RN success. The sample included 195 student records, which was the combined enrollment of the four semesters being studied, and was divided into two cohorts. The first cohort of 99 students from fall 2013 and spring 2014 did not have access to the CAQ software. The second cohort of 96 students from fall 2014 and spring 2015 had access to the CAQ software. The majority of students included were female, Caucasian, and in their early twenties at time of graduation.

**Procedures**

Prior to data collection, applications for Institutional Review Board (IRB) approval were submitted to the IRB committee at the University of Alabama as well as the IRB committee of the participating school. Exempt status was granted from both universities. Demographic data collected included race, gender, and age at graduation. Other data collected included (a) pre-nursing science GPA, (b) nursing GPA, (c) failure(s) in nursing courses, and (d) if another degree was held.
Demographic and nursing school data were collected from archived data through the student services department of the participating school of nursing. The course coordinator in the course that used the computer-adaptive software program collected PassPoint usage data. To protect the identification of students’ records, the course coordinator blinded all data prior to submitting to the investigator. The investigator then proceeded to numerically code all data received.

The students in the first two semesters, fall 2013 and spring 2014, of this study did not have access to a computer-adaptive software program. The selected product was implemented beginning in the fall 2014 semester and continued to be used during the spring 2015 semester. The product used in this study was PassPoint, a product of Wolters-Kluwer Health Publishing Company. Each final-semester nursing student was required by the school of nursing to purchase the product; individual informed consent or ethical releases were not collected. PassPoint was used during the final semester of the students’ curriculum as a preparation tool for the NCLEX-RN. Students had two proctored quizzes of 250 questions each, with a six hour time limit, at the beginning and end of the 12 week course. The quizzes were comprehensive of material learned during the entire course of study in the nursing program. Questions had varied formats including multiple choice, choose all that apply, matching, sequencing, and questions utilizing graphics. Since PassPoint is an adaptive software, questions were delivered based on individualized responses to questions at varying difficulty levels. At the end of each proctored testing session, the student was assigned, through the software program, an overall mastery level from one to eight as well as feedback for remediation needs based on the topics of the NCLEX-RN test plan. Students not reaching a mastery level of eight were encouraged by the course coordinator to
remediate through answering additional questions and taking additional practice quizzes. Any student, regardless of proctored exam mastery level, could access the program at any time for additional practice quizzes and remediation. Reasonable accommodations allowed for NCLEX-RN examinees were offered to the students but none chose to use them during the PassPoint testing sessions. Data including utilization of PassPoint for non-proctored practice quiz questions, average mastery level of practice questions answered, average mastery level on proctored exams, and first-time success on the NCLEX-RN were collected.

Demographic and nursing school data were collected from archived data through the student services department of the participating school of nursing. The course coordinator in the course that used the computer-adaptive software program collected PassPoint usage data. To protect the identification of students’ records, the course coordinator anonymized all data prior to submitting to the investigator. The investigator then proceeded to numerically code all data received.

Data Analysis

The independent variables for the study were the (a) access to a CAQ product, (b) average mastery level achieved on the proctored PassPoint exams, and (c) use of PassPoint for practice or remediation outside of the proctored exam times. The dependent variable of the study was the student’s pass/fail result on their first attempt of the NCLEX-RN. Following graduation and the first attempt of the NCLEX-RN of all students, data were collected indicating (a) if the student had access to PassPoint, (b) utilization of PassPoint for practice or remediation outside of proctored classroom exams, (c) average mastery level of proctored quizzes, and (d) if the student was successful on the first-time NCLEX-RN attempt.
Upon completion of data collection, information was entered into the Statistical Package for the Social Science ([SPSS], Version 19; IBM, 2010) software. Pearson’s chi-square analysis with an alpha of .05 and power of 0.8 were used to examine if access to the computer-adaptive testing product had an effect on NCLEX-RN success. Additional analysis was done to determine if the following student variables were correlated with NCLEX-RN success: (a) nursing GPA, (b) pre-nursing science GPA, (c) if an additional college degree was held, and (d) failure of any nursing courses. Due to the extremely small sample of NCLEX-RN failures in the second cohort, the investigator was unable to perform a regression analysis. Descriptive statistics were used to describe usage of PassPoint.

Summary

In summary, this chapter discussed the research design of this study, as well as the setting, population and sample, sample size, procedures, and data analysis. This study examined whether having access to and using a CAQ software affected NCLEX-RN results for undergraduate baccalaureate nursing students. Procedures and data analysis were described in detail.
CHAPTER 4

RESULTS

This chapter provides results of the study conducted to determine if using a CAQ program had a relationship to NCLEX-RN success. A description of the sample, demographic characteristics of the sample, and analysis of the equality of the two cohorts is included. A review of the research question and the hypotheses is provided. Results of the data analysis regarding access and usage as well as the explanation of analyses used are also included.

Research Question and Hypotheses

Research question: Does the use of a CAQ program as a tool for NCLEX-RN preparation during the final semester of a BSN curriculum predict first-time NCLEX-RN success?

Hypothesis 1: Students who utilized CAQ as a practice tool in addition to the proctored classroom quizzes will more likely successfully pass the NCLEX-RN on the first attempt.

Hypothesis 2: Students achieving an average mastery level of five or greater on two proctored CAQs will more likely successfully pass the NCLEX-RN on the first attempt.

Hypothesis 3: Having access to a CAQ software program will improve nursing students’ success rates on the first time NCLEX-RN attempt.

Description of the Sample

A purposive sample of de-identified nursing graduates’ records (N =195) from the fall and spring semesters during the academic years 2013-2014 and 2014-2015 were used for this
study. All 195 records were included in the study. Cohort one consisted of 99 records obtained from the fall 2013 and spring 2014 semesters. Cohort two consisted of 96 records obtained from the fall 2014 and spring 2015 semesters. Cohort one consisted of students who did not have access to PassPoint and cohort two consisted of students who had access to PassPoint. During the four semesters of data collection, there were no major curriculum or course faculty changes. There was no change in teaching strategies or course material delivery during the semesters of the study. Prior to the implementation of PassPoint with the second cohort, NCLEX-RN preparation tools were not part of the course.

The de-identified data were coded by the researcher using the basic Carnegie classification for institutions (Carnegie Foundation, n.d.). Codes for gender (1 = female, 2 = male), race (1 = Caucasian, 2 = African American, 3 = Asian, 4 = Hispanic, 5 = Other), additional degrees (1 = yes, 2 = no), nursing course failures (0 = no, 1 = yes), and NCLEX-RN (1 = pass, 2 = fail) were entered into Microsoft Excel software for Mac 2011 (Version 14.4.1: Microsoft, 2010) along with numerical data for age at graduation, pre-nursing science GPA, nursing GPA, and PassPoint mastery levels. Data were then copied into SPSS (Version 23; IBM, 2015).

Descriptive statistics of the demographic variables revealed a sample of 181 (92.8%) females and 14 (7.2%) males. One hundred and seventy-four (89.2%) students were Caucasian, nine (4.6%) were African American, seven (3.6%) were Hispanic, three (1.5%) were Asian, and two (1%) were listed as other. The average age at graduation of both cohorts was 24 years old. The majority, 149 (76.4%) did not have an additional college degree in another field, while 45 (23.6%) did hold an additional degree. Cohort one had 13 (13.1%) students with a nursing course
failure and cohort two had eight (8.3%) students with a nursing course failure. The average pre-
nursing science GPA was 3.17 for cohort one and 3.26 for cohort two. The average nursing GPA
for each cohort was 3.2. There was no attrition in this course over the four semesters of the
study.

Pearson’s chi-square analyses were used to determine if there was a significant difference
between the two cohorts for gender, race, nursing course failures, and other degrees held. The
results of the data analyses are reported in Tables 1 (gender), 2 (race), 3 (nursing course failures),
and 4 (other degrees held).

Table 1: Chi-square Tests for Gender

<table>
<thead>
<tr>
<th>Test</th>
<th>Value</th>
<th>df</th>
<th>Asymptotic Significance (2-sided)</th>
<th>Exact Sig. (2-sided)</th>
<th>Exact Sig. (1-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>2.576 (^a)</td>
<td>1</td>
<td>.109</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuity Correction (^b)</td>
<td>1.762</td>
<td>1</td>
<td>.184</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>2.660</td>
<td>1</td>
<td>.103</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fisher's Exact Test</td>
<td></td>
<td></td>
<td></td>
<td>.164</td>
<td>.091</td>
</tr>
<tr>
<td>Linear-by-Linear Association</td>
<td>2.562</td>
<td>1</td>
<td>.109</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>195</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. \(^a\) 0 cells (.0%) have expected count less than 5. The minimum expected count is 6.89. \(^b\) Computed only for a 2x2 table.
### Table 2: Chi-square Tests for Race

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>df</th>
<th>Asymptotic Significance (2-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>2.542</td>
<td>4</td>
<td>.637</td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>3.321</td>
<td>4</td>
<td>.506</td>
</tr>
<tr>
<td>Fisher's Exact Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear-by-Linear</td>
<td>.503</td>
<td>1</td>
<td>.478</td>
</tr>
<tr>
<td>Association</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>195</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. a. 8 cells (80.0%) have expected count less than 5. The minimum expected count is .98.

### Table 3: Chi-square Tests for Nursing Course Failures

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>df</th>
<th>Asymptotic Significance (2-sided)</th>
<th>Exact Sig. (2-sided)</th>
<th>Exact Sig. (1-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>1.168</td>
<td>1</td>
<td>.280</td>
<td>.357</td>
<td>.198</td>
</tr>
<tr>
<td>Continuity Correction b</td>
<td>.722</td>
<td>1</td>
<td>.396</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>1.179</td>
<td>1</td>
<td>.278</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fisher's Exact Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear-by-Linear</td>
<td>1.162</td>
<td>1</td>
<td>.281</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Association</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>195</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.34. b. Computed only for a 2x2 table
Table 4: Chi-square Tests for Other Degrees Held

<table>
<thead>
<tr>
<th>Test</th>
<th>Value</th>
<th>df</th>
<th>Asymptotic Significance (2-sided)</th>
<th>Exact Sig. (2-sided)</th>
<th>Exact Sig. (1-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>.308a</td>
<td>1</td>
<td>.579</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuity Correctionb</td>
<td>.150</td>
<td>1</td>
<td>.699</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>.309</td>
<td>1</td>
<td>.578</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fisher's Exact Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear-by-Linear Association</td>
<td>.307</td>
<td>1</td>
<td>.580</td>
<td>.616</td>
<td>.350</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>195</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 22.65. b. Computed only for a 2x2 table

These results indicated no significant difference between the two cohorts in regard to gender ($p = .109$), race ($p = .637$) nursing course failures ($p = .260$), or other degrees held ($p = .579$).

Levene’s tests for equality of variances were used to determine if there was homogeneity of the variances in relation to age, science GPA, and nursing GPA. Independent t-tests were used to analyze similarity of the averages of age, science GPA, and nursing GPA between the two cohorts. The results of the analyses are reported in Table 5.
The results indicated that the data was distributed normally and that there were no significant differences between the two cohorts in relation to age ($p = .657$), science GPA ($p = .111$), or nursing GPA ($p = .916$). From the results presented in Tables 1 through 7, the researcher assumed that the cohorts were similar and without statistically significant differences.
Data Analyses

Usage

To address hypothesis one and hypothesis two, PassPoint usage data were collected from cohort two \((n = 96)\) with 95 students who were successful on the NCLEX-RN and one student who was not successful. Descriptive statistics were used in analyzing PassPoint usage due to having only one unsuccessful student. For the 95 students who were successful on the NCLEX-RN, mastery levels on proctored test one ranged from 3.3-8.0, with a mean of 6.96; mastery levels on proctored test two ranged from 1.0-8.0 with a mean of 6.92. For the student who was unsuccessful on the NCLEX-RN, the mastery level of proctored test one was 8.0 and 5.2 for proctored test two. Practice quiz mastery level ranged from 0.0 to 6.7 with an average mastery level of 2.15 among successful students. The unsuccessful student did not utilize the quizzing software outside of the proctored tests thus no mastery level data were available for this student. Tables 6, 7, and 8 summarize PassPoint usage data for all variables for the second cohort. Having only one failure in the second cohort prevented data analysis of the usage data between students who passed and did not pass the NCLEX-RN. Although this analysis was not possible, the usage data collected highlights data that can be further researched.

The range of utilization of the computer adaptive software for practice outside of the proctored quizzes was wide and did not show significant patterns. Twenty two percent of the students in cohort two did not utilize the CAQ software outside of the proctored quizzes. Examining students who did answer practice questions outside of the proctored quizzes, 83% of students with a nursing GPA of 3.5-4.0, 83% of students with a nursing GPA of 3.0-3.49, and 73% of students with a nursing GPA of 2.5-2.99. Summary of this data can be found in Table 6.
Table 6: Cohort 2 Practice Question Data: Science GPA, Nursing GPA

<table>
<thead>
<tr>
<th># of Practice Questions (Range: 0-1392)</th>
<th>0</th>
<th>1-199</th>
<th>200-399</th>
<th>400-699</th>
<th>&gt;700</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science GPA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.5-4.0 (n=34)</td>
<td>13</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3.0-3.49 (n=34)</td>
<td>7</td>
<td>17</td>
<td>5</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>2.5-2.99 (n=23)</td>
<td>5</td>
<td>11</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2.0-2.49 (n=5)</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Nursing GPA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.5-4.0 (n=18)</td>
<td>8</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>3.0-3.49 (n=53)</td>
<td>11</td>
<td>24</td>
<td>9</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>2.5-2.99 (n=25)</td>
<td>7</td>
<td>10</td>
<td>2</td>
<td>6</td>
<td>0</td>
</tr>
</tbody>
</table>

Looking at the descriptive data for usage based on the individual variables identified in this study, a few observations can be made. Although the sample size was small, Hispanic and African American students answered more practice questions on average than Caucasian and Asian students. Hispanic students answered the most practice questions on average and had the highest average mastery level of practice questions, yet had the lowest average mastery level on both of the proctored quizzes. On proctored quiz one, average mastery level decreased as the science GPA and the nursing GPA decreased as well. For proctored quiz two, for students with a lower nursing GPA, the average mastery level, number of practice questions answered, and average mastery level of practice questions was less than those with higher nursing GPAs. Students with nursing course failures used fewer practice questions and had a lower average mastery level on quiz two compared to those students without nursing course failures. See Tables 7 and 8 for summary of descriptive usage data.
Table 7: Descriptive Data for Usage by Age, Gender, and Race

<table>
<thead>
<tr>
<th></th>
<th>Sample</th>
<th>Avg # practice questions</th>
<th>Avg mastery level of practice questions</th>
<th>Avg mastery Quiz 1</th>
<th>Avg mastery Quiz 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23-42</td>
<td>47</td>
<td>203</td>
<td>2.29</td>
<td>6.88</td>
<td>7.1</td>
</tr>
<tr>
<td>20-22</td>
<td>49</td>
<td>210</td>
<td>1.96</td>
<td>7.0</td>
<td>6.7</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>4</td>
<td>193</td>
<td>1.88</td>
<td>6.93</td>
<td>6.07</td>
</tr>
<tr>
<td>Females</td>
<td>92</td>
<td>207</td>
<td>2.13</td>
<td>6.98</td>
<td>6.94</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>4</td>
<td>614</td>
<td>3.3</td>
<td>5.55</td>
<td>6.2</td>
</tr>
<tr>
<td>Asian</td>
<td>1</td>
<td>20</td>
<td>2.0</td>
<td>8</td>
<td>7.7</td>
</tr>
<tr>
<td>African American</td>
<td>4</td>
<td>534</td>
<td>3.0</td>
<td>6.6</td>
<td>7.2</td>
</tr>
<tr>
<td>Caucasian</td>
<td>87</td>
<td>175</td>
<td>2.0</td>
<td>7.0</td>
<td>6.9</td>
</tr>
</tbody>
</table>

Table 8: Descriptive Data for Usage by Science GPA, Nursing GPA, and Nursing Course Failures

<table>
<thead>
<tr>
<th></th>
<th>Sample</th>
<th>Avg # practice questions</th>
<th>Avg mastery level of practice questions</th>
<th>Avg mastery Quiz 1</th>
<th>Avg mastery Quiz 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Science GPA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.5-4.0</td>
<td>34</td>
<td>222</td>
<td>2.15</td>
<td>7.2</td>
<td>7.17</td>
</tr>
<tr>
<td>3.0-3.49</td>
<td>34</td>
<td>157</td>
<td>1.87</td>
<td>6.93</td>
<td>6.65</td>
</tr>
<tr>
<td>2.5-2.99</td>
<td>23</td>
<td>243</td>
<td>2.47</td>
<td>6.79</td>
<td>6.77</td>
</tr>
<tr>
<td>2.0-2.49</td>
<td>5</td>
<td>273</td>
<td>2.06</td>
<td>6.4</td>
<td>7.36</td>
</tr>
<tr>
<td><strong>Nursing GPA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.5-4.0</td>
<td>18</td>
<td>247</td>
<td>2.37</td>
<td>7.68</td>
<td>7.62</td>
</tr>
<tr>
<td>3.0-3.49</td>
<td>53</td>
<td>208</td>
<td>2.17</td>
<td>6.87</td>
<td>6.86</td>
</tr>
<tr>
<td>2.5-2.99</td>
<td>25</td>
<td>174</td>
<td>1.86</td>
<td>6.68</td>
<td>6.48</td>
</tr>
<tr>
<td><strong>Nursing Course Failures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>8</td>
<td>179</td>
<td>2.25</td>
<td>7.0</td>
<td>6.16</td>
</tr>
<tr>
<td>No</td>
<td>88</td>
<td>209</td>
<td>2.11</td>
<td>6.97</td>
<td>6.97</td>
</tr>
</tbody>
</table>
Access and Usage

To satisfy hypothesis three, Chi-square analysis and logistic regression were conducted with degrees of freedom set at one to determine if the students who had access to and used the CAQ software, whether only during the proctored exams or during proctored exams and for additional practice, had an effect on success on the NCLEX-RN in comparison to the students who did not have access to the software (Table 9). Data from the chi-square analysis suggested that there was an association between having access to the CAQ software and success on the NCLEX-RN, \( \chi^2 = 14 \) \((df = 1, N = 195)\), \( p < .001 \) with a greater proportion of students who did not have access to the CAQ software being unsuccessful on NCLEX-RN (16.16%) compared to students with access to the CAQ software who were unsuccessful on the NCLEX-RN (1.05%).

Table 9: Chi-square Tests for Effect of CAQ on NCLEX-RN Success

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>df</th>
<th>Asymptotic Significance (2-sided)</th>
<th>Exact Sig. (2-sided)</th>
<th>Exact Sig. (1-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>14.001</td>
<td>1</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuity Correction</td>
<td>12.166</td>
<td>1</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>16.724</td>
<td>1</td>
<td>.000</td>
<td></td>
<td>.000</td>
</tr>
<tr>
<td>Fisher's Exact Test</td>
<td></td>
<td></td>
<td></td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Linear-by-Linear Association</td>
<td>13.930</td>
<td>1</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>195</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 8.37. b. Computed only for a 2x2 table

In the logistic regression, the predictive model, based solely upon the constant [NCLEX-RN success], significantly predicted 91.3% of students with CAQ software access having
NCLEX-RN success. The overall residual chi-square statistics of the variable not in the model [cohort groups] was $\chi^2 = 14.00$ with a $p$-value < .001 indicating that the addition of cohort groups variable would significantly affect its predictive power.

Logistic regression of the cohort groups showed $r^2 = .184$ indicating an 18% variance of the pass/fail can be explained by having or not having access to use of CAQ (Table 10). Therefore, the results suggested that there was an association between having access to the computer-adaptive quizzing software program success on the NCLEX-RN.

Table 10: Logistic regression of Use of CAQ Related to NCLEX-RN Success

<table>
<thead>
<tr>
<th>Step</th>
<th>-2 Log likelihood</th>
<th>Cox &amp; Snell R Square</th>
<th>Nagelkerke R Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>98.702$^{a}$</td>
<td>.082</td>
<td>.184</td>
</tr>
</tbody>
</table>

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

**Summary**

The investigator examined the data of 195 students from four semesters between August 2013 and May 2015 to determine if there was a relationship between using a computer-adaptive quizzing program and NCLEX-RN success. Analysis of the cohort data determined there to be no statistically significant differences between gender, race, age at graduation, pre-nursing science grade point average, nursing grade point average, and additional degrees earned. Based on data analysis in this study, results suggested that having access to and using a computer-adaptive quizzing software program may be helpful in passing the NCLEX-RN. Usage data was included in this chapter for descriptive purposes as analysis of usage of the computer-adaptive
quizzing software as well as average mastery of the proctored quizzes in relation to NCLEX-RN success in the second cohort was not possible as only one failure was recorded out of the 96 records. Discussion of these findings, limitations of the study, implications for nursing education, and recommendations for future research will be discussed in Chapter V.
CHAPTER V

DISCUSSION, IMPLICATIONS, AND RECOMMENDATIONS

This chapter includes a review of the research purpose, a review of the methodology used in this study, and a summary of the findings. Discussion of the results as they relate to the conceptual framework and in comparison to those of other researchers is included. Finally, implications for nursing education and recommendations for future research are provided.

Purpose Statement

The purpose of this study was to examine the relationship between use of a CAQ software program during the last semester of a BSN nursing curriculum and first time NCLEX-RN success. For students who had access to the CAQ tool, usage was defined as the number of practice questions answered, average mastery level of practice quizzes, and average mastery level of proctored quizzes. Student NCLEX-RN outcomes for those who had access to the CAQ tool and those who did not were compared to determine if access could be correlated with NCLEX-RN success.

Research question: Does the use of a CAQ program as a tool for NCLEX-RN preparation during the final semester of a BSN curriculum predict first-time NCLEX-RN success? Comparison of the two cohorts indicated no significant differences of race, gender, age, pre-nursing GPA, nursing GPA, prior degrees earned, and number of nursing course failures. Based on the data collected, 99% of the students who used PassPoint were successful on the first
attempt of the NCLEX-RN compared to 83% of students who did not use PassPoint during preparation for the NCLEX-RN.

_Hypothesis 1:_ Students who utilized CAQ as a practice tool in addition to the proctored classroom quizzes will be more likely to successfully pass the NCLEX-RN on the first attempt. Due to the small sample of one failure in the group of students using PassPoint, the amount of usage in relation to NCLEX-RN success was unable to be measured. As a note, the one failure did not use the CAQ software outside of the proctored quizzes.

_Hypothesis 2:_ Students achieving an average mastery level of five or greater on two proctored CAQs will be more likely to successfully pass the NCLEX-RN on the first attempt. As stated previously, due to having only one failure among the group using the CAQ software, usage data were unable to be measured. Of note, the student who was unsuccessful on the first attempt scored mastery levels of 8.0 and 5.2 on the two on the proctored quizzes respectively.

_Hypothesis 3:_ Having access to and using a CAQ software program will improve nursing students’ success rates on the first time NCLEX-RN attempt. The student cohort that had access to and used the CAQ software program had a significantly higher first-time success percentage (99%) compared to the cohort who did not use the CAQ software program (83%). No major instructional or curriculum changes were made over the four semesters included in the study.

**Review of the Methodology**

In this study, the investigator used a retrospective, descriptive, correlational study design. This design allowed the investigator to examine relationships between students’ access and usage of a computer-adaptive quizzing program and first-time NCLEX-RN success. A purposive
sample of de-identified, baccalaureate, final-semester nursing graduates’ records ($N = 195$) from the fall and spring semesters during the academic years 2013-2014 and 2014-2015 were used for this study. Cohort one ($n = 99$) did not have access to a computer-adaptive quizzing program while cohort 2 ($n = 96$) did have access. Data were analyzed to see if having access to a computer-adaptive quizzing program had an effect on NCLEX-RN success, if mastery level of proctored quizzes affected NCLEX-RN success, and if usage of the quizzing program affected NCLEX-RN results. Data analysis included Pearson’s chi-square analysis, Levene’s tests for equality of variances, and logistic regression.

**Summary of the Results**

The investigator examined the data of 195 final-semester, baccalaureate nursing students from four semesters between August 2013 and May 2015 to determine if there was a relationship between using a CAQ program and NCLEX-RN success. Analysis of the demographic data determined there to be no statistically significant differences between the two cohorts regarding gender, race, age at graduation, pre-nursing science GPA, nursing GPA, and additional degrees earned. Based on data analysis in this study, results suggested that having access to and using a CAQ software program during the final semester of a BSN program may be helpful in passing the NCLEX-RN. Analysis of mastery level and usage of the CAQ software in relation to NCLEX-RN success in the second cohort was not possible due to only one NCLEX-RN failure out of the 96 records.
Discussion of the Results

This section details the results of the statistical analysis for the research question and the three hypotheses. A discussion of the study in relation to the conceptual framework is also included. The results, as they relate to previous research, are discussed.

Phelan (2012) discussed that CAQs give the student and the instructor options to utilize formative assessment of competency and mastery and to individualize the learning process for each student. In this study, final-semester BSN students had access to a CAQ software program incorporating topics across the nursing curriculum. Each student was required to complete two proctored quizzes, had open use of the software for one calendar year, and was able to review for remediation any questions answered on prior practice exams.

CAQ follows the foundational principles of retrieval practice theory based on the concept that consistently accessing and applying topics previously learned and stored in the memory increases long-term retention (Dobson, 2013; Karpicke & Blunt, 2011; Pyc & Rawson, 2010; Roediger & Butler, 2011; Roediger & Karpicke, 2006a; Roediger & Karpicke, 2006b). This research supports the principles discussed by Roediger and Karpicke (2006b) of retrieval practice theory in that the repetition of answering questions in a quiz format, remediating material, and re-testing on a cyclical basis allows the learner to secure information in the long-term memory. Through a process of assigning cues to material stored in the long-term memory, the learner is able to later trigger the memory in order to retrieve previously learned material (Pyc & Rawson, 2010). The NCLEX-RN is given following graduation from an accredited school of nursing and tests information cumulative of material from the learner’s entire secondary education, thus requiring long-term retrieval of information. Having access to and using computer-adaptive
software gives the learner the opportunity to repeatedly practice retrieving information, applying it to questions in a format similar to the NCLEX-RN, and remediate concepts related to questions answered incorrectly. This process follows the foundational principles of retrieval practice to better prepare students for the NCLEX-RN.

With remediation allowing the student to become actively engaged in the learning process and putting the responsibility on the student, active learning and remediation become an integral component of content mastery (Bautsch, 2011; Wiles, 2015). Corrigan, Magaldi et al. (2014) discussed the importance of remediation in the learning process and Cox-Davenport and Phelan (2015) discussed the strong correlation of increased mastery of content in relation to increased usage of a CAQ program. Students in this study who had access to the CAQ program were required to review their proctored quizzes for remediation immediately following each quiz. Students who had access to the CAQ software program had significantly higher NCLEX-RN pass results compared to students who did not have access.

As discussed previously, usage and mastery level of proctored quizzes were not analyzed because the sample included only one failure on the NCLEX-RN in the group that had access to the CAQ software program. Many of the studies of NCLEX-RN success included in the review of literature stated that the limitation of minimal failures prevents results to be analyzed succinctly. The same result of few failures was true in this study.
Limitations

For the purposes of this research study, the following limitations were identified:

- The investigator did not have access to data describing students’ further preparation for the NCLEX-RN following graduation. Additional resources such as further studying, additional use of PassPoint following graduation, and NCLEX-RN preparation courses are all possible factors affecting the outcome of the NCLEX-RN. Likewise, individual student determination and focus cannot be measured.

- During the proctored PassPoint exams, a few students experienced technical issues such as loss of Internet connection, delayed question delivery, and having to exit and re-enter the exam. These factors could possibly lead to frustration and exam fatigue, causing skewed scores on the PassPoint exams. As the NCLEX-RN is also delivered via computer, technology issues are possible, but data related to these factors were not available.

- There was no recorded data about the length of time between graduation and the NCLEX-RN test dates of the students in the study. If the student waited a longer period of time, it could possibly negatively affect the outcome of the NCLEX-RN (Crow, Handley, Morrison, & Shelton, 2004; Daley et al., 2003; McGahee et al., 2010; Phelan, 2012, 2013; Yeom, 2013).

- Individual student academic ability is a limitation, which the investigator could only address through pre-nursing science GPA and nursing GPA. This factor cannot be controlled and serves as a limitation variable of success on the NCLEX-RN.
• Because of the small sample size of NCLEX-RN failures in cohort two, usage or mastery level data were not able to be compared based on NCLEX-RN outcome. Therefore limited findings could be analyzed from the study.

• The sample used for this study was composed from one, small, private university. The results may not generalizable to graduates of other programs.

**Implications for Nursing Education**

This research has several implications for nursing education. There is a limited body of knowledge in nursing education regarding CAQ as a tool for NCLEX-RN success. Research of trends and predictors of NCLEX-RN success is typically inconclusive or contradictory with one another. The findings of this study give nurse educators, as well as students, valuable information about a relatively new tool to help with preparation for the NCLEX-RN. The availability of a teaching tool and a preparation strategy that helps accomplish positive NCLEX-RN outcomes is a great advantage in nursing education.

With the *NLN Fair Testing Guidelines in Nursing Education* (2012) recommendations of decreasing high-stakes testing and progression policies based on testing outcomes, nurse educators must explore alternative strategies to facilitate NCLEX-RN success for students. CAQ software use allows simulation of the testing format as well as offers a source of feedback for both students and educators. If used throughout the nursing education curriculum, CAQ records can highlight at-risk students and serve as a remediation tool.

With the increase in nursing student enrollment, nurse educators must explore learning opportunities for self-learning and assessment for the student. CAQs allow for students to work
at an individualized pace, remediate material as needed, and retest the material with the opportunity of having differing questions. This sequencing of learning is important in that the CAQ software adapts to the student’s progression, or lack thereof, and gives continual updates of progression to both the student and faculty. This output can also guide the student and faculty instruction to explore possible misconceptions of various topics.

**Recommendations for Future Research**

Past research has studied predictor variables such as science GPA, ACT and SAT scores, reading scores, age, environmental stressors, high school education, and nursing GPA, with very few of these predictors showing a positive correlation with success on the NCLEX-RN (Frith et al., 2005; Simon et al., 2013). Research utilization of commercial high-stakes exams shows conflicting results on the predictive ability for NCLEX-RN success. There is limited research available about what strategies are the best for NCLEX-RN preparation, therefore, more studies are needed to determine strategies for promotion of student success, identification of at-risk students, and the most effective preparation strategies for diverse populations (Corrigan-Magaldi et al., 2014; Davenport, 2007; Jacobs et al., 2006; Lavin & Rosario-Sim, 2013; Romeo, 2013). Additional research on the ability to predict NCLEX-RN failure is also needed. However, with numbers of failure being small, based on national averages, data collection of predictors and preparation strategies is extremely challenging for this group of students (Crow et al., 2004, Daley et al., 2003, McGahee et al., 2010; Phelan, 2012, 2013; Yeom, 2013).

Computer-adaptive software programs are becoming more popular and readily available in nursing education, yet there is little research on the use of CAQ as a tool for NCLEX-RN
preparation. With the ability to implement CAQ as a teaching strategy for educators, a self-learning and assessment strategy for students, an identifier of high-risk students, and a remediation tool, further research is needed to examine if utilizing a computer-adaptive model is an effective tool for NCLEX-RN success.

**Study Summary**

With NCLEX-RN results being a significant measure of program quality (Giddens, 2009; McGahee, Gramling, & Reid, 2010; Phelan, 2012; Simon et al., 2013), the importance of preparing nursing students for success is essential. With little conclusive evidence of predictors for NCLEX-RN success, it is imperative that nurse educators explore preparation options that meet student needs and continually challenge learning and ability. Although CAT has existed for many decades in other disciplines, the phenomenon is relatively new in nursing education. The NCSBN began administering the exam in a computer-adaptive format in 1994, (NCSBN, 2015) allowing graduates’ knowledge to be better assessed in comparison to paper-and-pencil testing. This transition to CAT pushed nurse educators to reexamine how students were being prepared. Schools of nursing began computer testing in the 1990’s but the adaptive technology was missing.

In 2009, computer-adaptive technology within the curriculum was introduced to nurse educators in the United States (Wolters Kluwer, 2014). This gave nurse educators and graduates an opportunity to gauge knowledge and mastery on an individualized basis. CAQ, the process of studying, testing, re-studying, and testing follows the foundation of retrieval practice theory. Following initial exposure to material, students can test using a CAQ platform, re-study
identified subjects of weakness, and retest on a continual rotation until competency is achieved. This gives graduates the opportunity to tailor their learning to specific needs, on a personal schedule, and measure mastery in comparison to classmates and others nationally. With CAQ being very new to nursing education, the review of literature exposed a gap in the literature. Research about predictors of nursing school success, NCLEX-RN success, and various other teaching strategies were numerous, but very little research has been published about using CAQ as a preparation tool for NCLEX-RN success. This gap in the literature provides opportunities for contributions to the body of knowledge for nursing education by conducting additional research.

This study explored use of the PassPoint software program as a tool for NCLEX-RN preparation. There was no significant difference among students in regard to gender, age at graduation, pre-nursing and nursing GPA, nursing course failures, or if other degrees were held. There was no change in curriculum, course materials, or course instructors during the four semesters of this study other than the implementation of the PassPoint software in cohort two. Results showed that students who had access to and used a CAQ software program during the final semester of a BSN curriculum had 99% success on the NCLEX-RN compared to 83.4% of students who did not have access. These results provide the foundation for further research on best practices for implementation of CAQ software into nursing curriculum as well as whether or not individual usage patterns of the software have an impact on NCLEX-RN success.
REFERENCES


