2015

Assessment Practices at an Associate Degree Nursing Program

Tracey Jane Siegel
Walden University

Walden University

COLLEGE OF EDUCATION

This is to certify that the doctoral study by

Tracey Siegel

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Sydney Parent, Committee Chairperson, Education Faculty
Dr. Dustin Hebert, Committee Member, Education Faculty
Dr. Mary Howe, University Reviewer, Education Faculty

Chief Academic Officer

Eric Riedel, Ph.D.

Walden University
2015

Abstract

Assessment Practices at an Associate Degree Nursing Program

by

Tracey J. Siegel

MSN, Monmouth University, 2000

BSN, Rutgers University, 1980

Doctoral Study Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Education

Walden University

April 2015

Abstract

Nursing programs have traditionally used teacher-developed multiple-choice (MCQ) examinations to prepare students for licensure. Researchers have determined that poorly constructed MCQ tests used as formative and summative evaluations may penalize nursing students and impact progression and retention in nursing programs. The purpose of this exploratory case study was to examine issues related to the use of teacher-developed MCQ examinations as the only method of student assessment in the theory component of nursing courses. The National League for Nursing Core Competencies for Nurse Educators and the revised Bloom's Taxonomy were used as the conceptual frameworks for this study. The Director of the Nursing Program and 9 faculty members participated. Data were collected from a review of documents, 2 focus groups, faculty-maintained diaries, and an interview. During data analysis, categories were identified and themes emerged, revealing the key findings. Using a single method alone to assess student learning limited the opportunity for formative assessment, the ability to assess higher order thinking, and the development of metacognition on the part of students. To assist faculty in creating assessments of student learning that would address these themes, a 3-day faculty professional development project followed by 4 monthly lunch and learn sessions was designed. Providing additional faculty development in assessment methods may promote positive social change as it may ultimately increase the retention of qualified students to meet the demand for registered nurses within the community.

Assessment Practices at an Associate Degree Nursing Program

by

Tracey J. Siegel

MSN, Monmouth University, 2000

BSN, Rutgers University, 1980

Doctoral Study Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Education

Walden University

April 2015

Dedication

"There's something like a line of gold thread running through a man's words when he talks to his daughter, and gradually over the years it gets to be long enough for you to pick up in your hands and weave into a cloth that feels like love itself"

John Gregory Brown (Decorations, 1994)

To my father, George David Graul who always believed that I could be anything I wanted to be and do anything I wanted to do. He was so proud of me when I graduated from nursing school, yet always told me that I should have been a doctor. Well Dad, that is finally happening and I wish you could have lived to see it.

To my mother, Theresa Veronica Griffen Graul; you were my rock, my guide, and my very best friend. Thanks for your support and encouragement during this entire process. I am sorry that you did not live to see this project completed. Even at the end you would ask me "how is school going?" What an amazing lady you were!

And last, but definitely not least, to my family. Bernie, I do not know how I could have completed this without your support and unconditional love. You kept telling me not to give up and to keep going. Bernard, your daily encouragement and sense of humor about this entire process kept me going. Matthew, every day you would ask "Hey mom, how's the paper coming?" You have no idea how much that meant to me. All of you are my life.

Acknowledgments

I would like to publically thank and acknowledge the following individuals:

- Sydney Parent, my capstone chairperson who has amazing insight but more importantly a wonderful sense of humor. Thanks for your encouragement and keeping me grounded.
- Nancy Berger, my mentor, my colleague, but most of all my friend. At least once a week you would say, "Can I do anything at all to help you?" I know that these were never empty words. I am blessed.
- The faculty of my nursing program for assisting me so readily with this capstone project. You were open and honest but more importantly, supportive. Thank you for being wonderful colleagues and friends.

Table of Contents

List of Tables

Section 1: The Problem

Introduction

In September 2007, a new associate degree nursing program was established at a suburban community college located in one of the eastern Mid-Atlantic States. Based upon the recommendations of a curriculum expert, the faculty decided to use teacher developed multiple-choice question (MCQ) examinations as the only method to assess student knowledge and assign grades in the theory portion of each course. Over the past 7 years, however, based upon student feedback and attrition, faculty members have expressed concerns about their ability to design fair, reliable, and valid multiple-choice question examinations. Test item analyses, test blueprints, student course end evaluations, and one grievance to college administration also support this problem. In curriculum meetings, faculty have reported "feeling stressed and upset" during test construction and administration, and often "dread" test reviews (Program Coordinator, personal communication, January 15, 2011).

To prepare students for licensure, nursing programs have traditionally used teacher-developed MCQ examinations for both formative and summative assessment (Killingsworth, 2013; McDonald, 2013; Oermann & Gaberson, 2009; Oermann, Saewert, Charasika, & Yarbrough, 2009; Redmond, Hartigan-Rogers, & Cobbett, 2012; Walloch, 2006). It is the opinion of some nursing education experts that the best way to prepare students for the National Council Licensing Examination (NCLEX-RN) is to use the same format as the NCLEX (Yoho, 2010).

Nursing is unique in that professional licensing standards require graduates to pass this comprehensive, primarily MCQ computer-adapted examination. The initial passing rates of new graduates determine accreditation status and are used as a program learning outcome for all undergraduate nursing schools (Accreditation Commission for Education in Nursing Standards, ACEN, 2013). Despite research supporting the use of MCQ that test at the application and analysis levels for high-stakes testing (Brady, 2005; Draper, 2009; Leung, Mok, & Wong, 2008; Wendt & Harmes, 2009; Williams, 2006), some nursing education experts have recently called for radical change in how student learning outcomes are assessed to ensure that all domains of learning are included and that assessments are fair to diverse learners (Benner, Sutphen, Leonard, & Day, 2010; Giddens, 2009; Oberman & Gaberson, 2009).

The faculty at this community college were determined to use evidence-based assessment and teaching methods yet were worried that changing current practice will undermine initial NCLEX-RN pass rates and impact accreditation status. Faculty also were concerned that the addition of other assessment methods for grading purposes may allow students to attain passing grades without ever passing an examination. The goal of this study was to explore faculty and student issues related to using only one method to assess student learning and assign grades in nursing theory courses.

Definition of the Problem

The program had excellent NCLEX-RN pass rates for the first four graduating classes (97%, 94%, 92%, and 93.6%), and has achieved full initial accreditation from the National League for Nursing Accreditation Commission (NLNAC); however, there has

been significant student attrition. Students are allowed 3 years to complete this 2-year program, and are afforded the opportunity to return one time after a failing grade. Therefore, attrition rates are calculated for students completing the program within the 2 or 3-year timeframes. The attrition rates for the program, since its inception, are outlined in Table 1 and are similar to what has been reported in the literature (e.g., Rees, 2006). However, these reported rates are over the course of entire programs, not one semester. Although college administration has not as yet asked for a full accounting of this high attrition rate, faculty are concerned. Anecdotally, based upon student course end evaluations, conversations, and cursory reviews of examinations, some part of this attrition rate might be attributed to issues with test construction.

Table 1

Attrition Rates by Course and Class

Course	Class of 2009	Class of 2010	Class of 2011	Class of 2012
Nursing 1	4%	4%	4%	13%
Nursing 2	25%	32%	25%	29%
Nursing 3	27%	8%	22%	8%
Nursing 4	4.5%	3%	1%	3%

Course end evaluations that were previously shared by program administration have noted that students felt unprepared for nursing multiple-choice questions and some students perceived the examinations as unfair. Included on the course evaluations for

prior semesters were many student comments about the faculty constructed examinations. Student comments included:

- "Faculty should spend more time writing test questions"
- "In the future there should be more focus on test questions"
- "Is there any way our grades could be based on more than our test grades?"

Faculty admitted that some examinations had issues with grammar, stem construction, and poorly designed distracters. Students must pass both the theory and clinical components of each course to progress in the program. All final grades in the theory portion of each course are based solely upon examination grades; therefore, issues with this assessment method could potentially penalize students (Tarrant, Knierim, Hayes, & Ware, 2006; Tarrant & Ware, 2008) and prevent continuation in the nursing program. Students have been quite verbal during faculty conducted test reviews and have complained to the nursing program administration about perceived issues with the MCQ examinations. Each testing incident has been viewed as a single entity without evaluating the overall process of student assessment. It is unknown why faculty have maintained that the best way to assess student learning and assign grades in the theory portion of each nursing course is only by the administration of teacher-developed multiple-choice tests.

This nursing program has a large percentage of foreign educated and limited English proficiency adult learners. Nationally, these students are often penalized by poorly designed MCQ examinations (Bosher & Pharris, 2009). Diverse students have unique and varied learning needs and styles (Brown, 2008; Starr, 2009); however, many

nurse educators unconsciously teach the way they were taught and use assessment methods designed for the dominant culture (Ackerman-Barger, 2010).

Depending on their educational background, nursing faculty may have received little to no formal training in assessment practices (Walloch, 2006). Foreign educated students and those with limited English proficiency often have higher attrition rates (Jeffreys, 2012). In 2013, nearly 60% of the student body in this community college was diverse or nontraditional according to Jeffreys's criteria (College Demographics, 2013). The nursing program has no hard data about the attrition rates of nontraditional students. Anecdotally, observations and interviews with failing students support the contention that the majority of the students who failed nursing courses were nontraditional, educated in foreign schools, or spoke English as an additional language (Program Coordinator, personal communication, January 24, 2012). Researchers have clearly demonstrated poorly constructed MCQ examinations penalize not only diverse and nontraditional students but all students, and impact progression and retention in nursing programs (Bosher & Pharris, 2009; Clifton & Schriner, 2010; Dowling, 2005; Tarrant, Knierim, Hayes, & Ware, 2006; Tarrant & Ware, 2012).

Rationale

Evidence of the Problem at the Local Level

The associate degree nursing program is part of a community college located in a suburban area in one of the eastern Mid-Atlantic States. It serves more than 14,600 students representing a diverse array of cultures, backgrounds, ages, and educational needs (College Demographics, 2013). The program was established in 2006 and it has a

unique design. The community college owns the program; however, the faculty are employees of a local hospital. This joint program benefited both parties. Hospital administration wanted to close its Diploma School of Nursing and the community college was losing its affiliation with a different nursing program. All faculty were seasoned educators, however, the majority had no experience teaching in a community college setting. The diploma program was 3 years in length; therefore, a new curriculum design was needed to fit the community college's four semester model. There are rigorous admission guidelines and 90 students are admitted annually each September. In 2009, the program attained full initial accreditation from the NLNAC, and in 2012, a full 8-year accreditation from the state Board of Nursing.

In 2013, there were 13 full-time faculty who were all women between the ages of 40 and 64, and only three were culturally diverse (African-American, Hispanic, and Asian Indian). All faculty had a master's degree in nursing, however, only six were prepared as educators at the graduate level. The remaining faculty were prepared as clinical specialists, nurse practitioners, or administrators. Four of the faculty who were not prepared as nurse educators had attained the prestigious Certified Nurse Educator (CNE) credential from the National League for Nursing, which acknowledges expertise in eight core educational competencies.

Most of the faculty did not have any formal training in student assessment or evaluation on the graduate level (Nursing Program Board of Nursing Self Study, 2012). However, all faculty had been actively involved with professional development, and had attended numerous workshops and conferences about the assessment and evaluation of

student learning outcomes. Over the past 6 years, all faculty participated in at least four professional development courses about MCQ item writing and test construction. These include the following by name:

- The Ultimate Test: Developing Valid and Reliable Exams to Assess Higher-Order Thinking (New Jersey League for Nursing, 2005),
- Be an Innovative Nurse Educator: Strategies that Promote Classroom and Clinical Learning (Herman, 2010),
- Regional NCLEX Test Writing Workshop (National Council of State Boards of Nursing, 2010),
- Test Construction for Nurse Educators (Ortelli, 2011).

Improving the quality of teacher-developed tests has been a priority of administration and the course coordinators. At every curriculum meeting since January 2009, the program coordinator has stressed test blueprinting, item analysis, form and style, and the development of valid and reliable test banks (Program Coordinator, personal communication, January 24, 2012). Despite this emphasis on improving tests, in 2011 and 2012 the evaluations from the second and third semester courses contained pages of student comments about the MCQ tests. For example:

- “I hope that, in the future, there will be more focus on test questions and how to answer them.”
- “Is there any way that our grades weren't solely based on our exams? We place a great deal of time and energy in our clinicals as in our lab sessions with the presentations and test outs. Couldn't they be worth some percentage

instead of them just being a Pass/Fail? Those areas are just as important as the exams because we are actually applying what we've learned."

- "There were numerous typos and problems with the wording of this semester's tests as well as questions with multiple correct answers…it isn't fair that we are penalized."
- "Some of the faculty's questions were difficult to understand and didn't reflect what was covered in class."

Using teacher-developed examinations as the only method to assess students and to assign grades has impacted the faculty. Many faculty were originally strong proponents, believing that students would benefit from repeated testing using the same format as the NCLEX-RN. However, after many highly qualified and clinically proficient students failed the theory content of the second and third semester courses, some faculty began to question this practice (Systematic Plan for Evaluation Committee Minutes, June 2011). Many students who failed were nontraditional, culturally diverse, or had limited English proficiency. To attempt to address student issues with MCQ examinations, some faculty members have spent countless hours tutoring students and reviewing tests.

Faculty use multiple methods to assign clinical grades including observation, reflective journals, patient care plans, and clinical simulation. Students receive a midterm formative assessment which provides them with clear feedback for areas needing further development and improvement. This is followed 6 weeks later with a final summative clinical evaluation.

There are few opportunities for formative assessment in the theory component of each course. For example, during the second semester medical surgical nursing course, students take four MCQ examinations consisting of 50 questions each and a final examination which is considered summative and consists of 75 questions; 50 questions from new content and 25 questions reflecting all course content. It could, therefore, be argued that each test is a formative appraisal of student learning, however, each test is summative for an area of medical surgical nursing. It is impossible for students to revisit areas of weakness or confusion; once tested students move immediately into another area of medical surgical nursing. This means that students who attain a failing grade have little opportunity to demonstrate understanding of tested content.

It could also be argued that students are given formative feedback during faculty led *test reviews* following each test. However, test reviews were conducted in a large lecture hall with up to 90 students in attendance, and there was little consistency among faculty about the best ways to provide feedback. Some faculty provided students the correct answer and gave a rationale which consists of a page number in the textbook explaining the correct response. Other faculty spent a little more time explaining the rationale; while, still others spent a large amount of time helping students understand the question stem, each distracter, and why each choice is correct or incorrect. Following the test reviews, students could also opt to schedule an individual review of the test with a faculty member.

Students were also advised to practice taking MCQ tests by using resources from textbooks and a comprehensive testing and review package purchased by students each

semester. Students were expected to use these resources as a self-assessment and to practice for the teacher-developed MCQ examinations; however, they did not receive any formal faculty feedback. Roedinger and March (2005) stressed that MCQ assessments expose students to incorrect answers and without feedback from faculty, students may not understand why incorrect answers are wrong.

Evidence of the Problem from the Professional Literature

Assessment experts (Knight, 2006; Nicol & Macfarlane-Dick, 2006; Stiggins & Du Four, 2009) emphasized the importance of multiple types of student assessment, and the use of both formative and summative methods to determine if learning outcomes are met. Multiple-choice and other selected-response question tests are considered adequate when determining lower level objectives but are not designed to capture the ability to create and synthesize, nor does this method encourage metacognition (Anderson & Krathwohl, 2001; Knight, 2006; Nilson, 2010).

Constructing reliable and valid assessments of student learning has long been an issue in nursing education. Many nurse educators were not prepared as teachers, but as clinicians or managers and therefore have little experience with MCQ item writing, test construction, or other forms of student assessment. Nursing education researchers have outlined numerous variables impacting MCQ test construction including poorly constructed textbook test banks, faculty inexperience, and lack of educational preparation (Clifton & Schriner, 2010; Killingsworth, 2013; Mc Donald, 2013; Oermann & Gaberson, 2009; Su, Osisek, Montgomery, & Pellar, 2009); however, there is little in the

literature which discusses the issues regarding using teacher-developed MCQ tests as the only method for assessment of students' knowledge of theory content.

Definitions

The following terms and phrases are defined as used in this study.

Associate degree nursing program. It is a two year academic program located primarily in community colleges which prepares students for nursing licensure. The first associate degree programs were established in 1957 in response to a profound nursing shortage and a call to educate nurses in academic, not hospital settings (Starr, 2010).

Formative assessment. Following an extensive review of the literature, Baroudi (2007) defined formative assessment as "activities used by the teacher to determine a student's level of knowledge and understanding for the purpose of providing students with feedback and planning future instruction. The feedback and future instruction may be concerned with remediation or the provision of further learning opportunities" (p. 39).

Multiple-choice questions. A form of selected-response problem designed as a question or incomplete statement and four possible options to answer the question or complete the statement (Oermann & Gaberson, 2009). In nursing education, the question typically contains a stem which presents a clinical scenario or question followed by four plausible options of one correct answer (key) and three distracters.

Teacher developed tests and examinations. A paper and pencil assessment designed by teachers for a specific group of students and administered under specific conditions within the classroom setting (Frey &Schmitt, 2007).

Significance

There is currently a global nursing shortage and a documented need for a more diverse nursing workforce (Fulcher & Mullin, 2011; Institute of Medicine, 2011; National League for Nursing, 2010; Rossiter, 2010). Suskie (2009) explained that assessments need to be accurate, unbiased, and fair to all students. Therefore, it is imperative that nursing programs use unbiased, objective, and impartial assessments to afford all students an equal opportunity for success.

Retaining qualified students in nursing programs is important to meet the local and global need for registered nurses. It is estimated that, by 2018, over 1 million qualified nurses will be needed to resupply the number of registered nurses to meet demand (Bureau of Labor Statistics, 2010). In addition, the population of adults over the age of 85 is growing faster than the population of individuals under the age of 45 (HRSA, 2010). The majority of nurses employed in geriatric long term care settings or in rural and underserved communities were initially educated in community colleges (Starr, 2010). In 2010, over 45% of practicing registered nurses were initially educated in associate degree programs (HRSA, 2010); and these programs supply the largest number of culturally and ethnically diverse nurses.

Health care organization leaders surmised that communication among healthcare professionals, the quality of patient care, and overall safety will be enhanced by a more diverse nursing workforce (Institute of Medicine, 2010; National League for Nursing, 2010). Despite this, the enrollment and retention of ethnically diverse students remains low; and the number of ethnic minorities represented in the nursing profession has

changed little over the past 20 years (National League for Nursing, 2009). Nurses need to be expert critical thinkers, demonstrate clinical judgment, and be reflective practitioners (Benner et al., 2010), and these characteristics may be difficult to evaluate by using only one method of assessment. The purpose of this study is to describe the faculty and student issues related to using teacher-developed MCQ examinations as the only method to assess students' knowledge of nursing theory and to assign grades.

Guiding/Research Question

While MCQ examinations can test knowledge at the applying and analyzing levels of the revised Bloom's taxonomy, registered nurses need more than cognitive knowledge; they need to be expert critical thinkers who have the ability to use clinical reasoning and sound judgment in today's complex patient care situations. The majority of nursing programs continue to use teacher-developed MCQ examinations to assess student learning despite calls for the use of innovative and evidence-based assessment methods.

The following question guided this research: What are the issues related to the use of teacher-developed MCQ examinations as the only method of student assessment and evaluation in the theory component of courses at a suburban associate degree nursing program?

Subquestions include:

1. How has using only teacher-developed MCQ examinations impacted the teaching/learning process?

2. How has using only teacher-developed MCQ examinations impacted the teacher/ student relationship?
3. How are students assessed formatively in the theory component of the nursing courses?

Little is known about how this assessment method impacts students, affects the practice of the faculty, or impacts the learning outcomes of the nursing program. Therefore, a case study was the most appropriate design for this study as it is an attempt to explore a problem on a local level. A case study is an in-depth analysis or description of an individual, group, situation, or event with an attempt to understand a complex issue through the eyes of those involved (Yin, 2009).

Review of the Literature

To understand this issue, I conducted an extensive review of both the nursing and general educational literature using the following databases from the Walden University Library: Academic Research Complete, Educational Resources Information Center (ERIC), Educational Research Complete, Cumulative Index to Nursing and Allied Health Literature (CINAHL) and Medline Simultaneous, Walden Thoreau, and Google Scholar. Of these databases, CINAHL and Google Scholar yielded the most articles about nursing education. The search parameters included only peer reviewed journals and included research and other types of articles such as meta-analyses and reviews. The original search for articles used data from 2008 to the present; however, to determine the origins of student assessment methods in nursing education and the theoretical framework, the search timeline was expanded. The following keywords and combinations were used:

teacher made tests, *teacher made multiple choice question tests*, *teacher made tests* and *nursing education*, *assessment of nursing student learning outcomes*, *student assessment* and *higher education*, *nursing faculty perceptions of student assessment*, *nursing student perceptions of assessment methods*, and *nursing faculty* and *student assessment methods*. Additionally, the references from three textbooks about assessment and evaluation in nursing education (Billings & Halstead, 2009; Oermann & Gaberson, 2009; Mc Donald, 2013) and from all pertinent peer review articles were reviewed.

Conceptual Frameworks

In 2001 the National League for Nursing convened a task force to perform a comprehensive and exhaustive review of the literature to determine the core competencies and best practices of nurse educators (Ortelli, 2006). The task force comprised educational experts from nursing, medicine, and higher education. Following this literature review, a practice analysis survey was developed and administered electronically to a representative, randomly selected sample of academic nurse educators throughout the United States. Based upon this practice analysis, the National League for Nursing's Core Competencies of Nursing Educators (2005) was designed. This document defined eight competencies critical to the role of an effective nurse educator.

According to these competencies, the nurse educator should:

1. facilitate learning,
2. facilitate learner development and socialization,
3. use assessment and evaluation strategies,
4. participate in curriculum design and evaluation of program outcomes,

5. function as a change agent and leader,
6. pursue continuous quality improvement in the nurse educator role,
7. engage in scholarship and,
8. function within the educational environment. (Core Competencies, pp. 1-8, 2005)

Nursing education leaders (Kalb, 2008; Utley, 2011) suggested that these competencies provided clear guidelines for practice and give nurse educators a method for reflection and self-evaluation. Competency 3 stresses that the nurse educator use a variety of evidence-based assessment methods to enhance teaching and learning in classroom and clinical settings and will be one of the conceptual frameworks guiding this study.

The second conceptual framework is the revised Bloom's taxonomy (Anderson & Krathwohl, 2001). Bloom's (1956) classic work described methods to quantify the three domains of learning, the cognitive, the psychomotor, and the affective. According to Krathwohl (2002), it was designed to give educators a common language when writing objectives and determining student learning needs. The cognitive domain consisted of six categories: "knowledge, comprehension, application, analysis, synthesis, and evaluation", and these were ordered "from simple to complex and from concrete to abstract" (p. 212). Anderson and Krathwohl (2001) revised this taxonomy to include different types of knowledge in a two-dimensional format. These knowledge types include the factual, conceptual, procedural, and metacognitive.

This taxonomy has recently been incorporated into nursing education. In five articles, Su and associates (2011, 2009, 2007, 2005, & 2004) have described methods to use this taxonomy when writing objectives and designing appropriate assessments based upon knowledge types and cognitive processes. The National Council of State Boards of Nursing uses this model as well when designing all MCQ for the NCLEX-RN examination. To pass this examination, candidates must answer factual, conceptual, and procedural questions written at the revised applying and analyzing levels (NCLEX-RN Test Plan, 2013).

There is little in the literature describing issues with using only one form of student assessment to evaluate learning or assign grades. Therefore, in this literature review, I focused on the following areas (a) MCQ tests as a method to assess student learning, (b) the use of MCQ examinations as an assessment method in nursing education, (c) the impact of flawed multiple-choice question tests, and (d) the use of multiple-choice question examinations with students who have limited English proficiency.

Multiple-choice Question Tests as a Method to Assess Student Learning

Multiple-choice question tests are frequently used as a method to assess student learning. These tests are considered objective, and easy to prepare, administer, and grade. Multiple-choice examinations allow for timely feedback, may reduce student anxiety, and if properly constructed are an excellent method to assess most cognitive levels. Multiple choice question examinations also facilitate the testing of large numbers of students and discourage cheating by allowing numerous versions of the same test to be

administered at the same time (Mc Donald, 2013; Nilson, 2010, Oermann & Gaberson, 2009; Roedinger & Marsh, 2005).

Haladya and Dowling (1989) developed a taxonomy which is considered the gold standard of MCQ item writing rules. This model was based upon extensive research and it delineated best practices for MCQ item writing. The taxonomy outlined 46 rules for designing fair, unbiased, reliable, and valid MCQ examinations. Haladyna, Dowling, and Rodriquez (2002) reviewed the literature again and revised the number of rules to 31 that were supported by additional research.

In nursing education, Mc Donald (2013) integrated Haladya and Dowling's taxonomy into guidelines, templates, and style guides. Mc Donald's guidelines included 32 recommendations for designing MCQ examinations and are used as a model for writing standardized examinations and the NCLEX-RN. Mc Donald gave nurse educators clear methods to target cognitive levels and outlined sample stems or questions to assist nursing faculty when developing tests.

Numerous research studies (Anderson & Krathwohl, 2001; Clifton & Schriner, 2010; Knight, 2006; Oermann & Gaberson, 2009; Su, Osisek, Montgomery, & Pellar, 2009) in education and health-related disciplines examined best practices for MCQ item writing. The psychometric properties of these examinations, comparisons to constructed response questions such as short answer, fill in the blanks, and essays, and methods to determine reliability and validity have been studied extensively (Rodriquez, 2005; Roedinger & Marsh, 2005). The prevailing belief was that MCQ cannot adequately assess the application of knowledge, a critical aspect in health-related disciplines.

This contention was challenged by Palmer and Devitt (2007) who analyzed modified essay and MCQ formats used for summative assessment in an Australian medical school. Three final examinations containing 50 MCQ and 139 modified essay questions were reviewed. The authors found that, despite purporting to assess the ability of students to apply or analyze content, over 50% of the modified essay questions actually tested factual recall, while the MCQ frequently assessed students at higher cognitive levels.

Using well-constructed MCQ examinations to assess the application of course content was the focus of research conducted by Stupans (2006). Using a retrospective quantitative design, pharmacology examinations from a 10 year period were reviewed. Cognitive levels of questions and statistical analyses supported the use of MCQ tests to adequately assess the application, analysis, and evaluation of pharmacology concepts. The author did caution that questions needed to be well constructed and noted frequent variability of cognitive level determinations by faculty.

In a more recent study also conducted in a pharmacology program, Tiemeier, Stacy and Burke (2011) determined that well-constructed MCQ examinations did assess students at higher cognitive levels. Students were given tests that contained 40% recall questions, 40% application questions and 20% analysis questions (as determined by the researchers) in either a MCQ or constructed response format. Using correlational analysis, the scatterplots for both types of questions were similar and the constructed response questions did not always assess higher cognitive levels as hypothesized. Again,

the authors cautioned that MCQ needed to be well designed and conform to item writing standards.

In a correlational study designed to determine if MCQ tests assessed similar knowledge levels as constructed response questions in an accounting course, Bible, Simkin, and Kuechler (2008) concluded that both types of questions measure essentially the same knowledge levels. The researchers analyzed the tests of 238 students enrolled in an intermediate accounting course over four semesters, and by using correlational statistics ($r2 = 66$) determined that MCQ are an adequate replacement for essay questions for testing higher level accounting concepts.

Test enhanced learning or the testing effect has been researched extensively. As Roediger and Karpicke (2006) explained, "Testing has a powerful positive effect on future retention. If students are tested on material and successfully recall or recognize it, they will remember it better in the future than if they had not been tested" (p.249). The Memory Laboratory of Washington University at St. Louis has conducted numerous studies into test enhanced learning and determined that this phenomenon occurs with all types of question formats. Multiple-choice question tests can boost retention especially in circumstances when students are required to analyze options rather than simply recognize or recall facts (Butler & Roediger, 2008; Marsh, Roediger, Bjork, & Bjork, 2007). The researchers cautioned that repeated use of MCQ exposes students to incorrect responses and may lead to the acquisition of false knowledge; however, they opined that the overall positive testing effect outweighs this potential problem.

While MCQ tests can test the ability to apply knowledge and analyze concepts, numerous authors note that this type of assessment cannot measure synthesis or metacognition (Krathwohl, 2002; Rodriquez, 2005). This contention was supported by Kim et al. (2009) who used a mixed methods design to compare constructed response and selected response answers for simulated patient care experiences in medical students enrolled in an on-line therapeutic communication course. Students observed a clinical scenario which required a response to a simulated patient's questions or requests for teaching. Following the scenario, students were required to write the answer to this question, "In your own words, what would you say next to the patient?" (p. 535). The students were then asked to select an appropriate response from four options. The written responses were compared to the selected-response choices and the results were surprising. The open-ended responses were medical terminology centered and difficult for patients to understand, while the selected responses were based upon recognition of the appropriate methods for therapeutic communication.

The researchers (Kim et al, 2009) concluded that the students recognized the correct response when listed, but tended to use incorrect communication techniques when they were required to construct a response. This was an important outcome as it reflected what the students would actually say in clinical situations. It also allowed the faculty to more fully appreciate and understand the students' thought processes and to provide feedback.

Another concern about the extensive use of MCQ test is teaching to the test, particularly in standardized high stakes examinations. Using a metasynthesis approach,

49 qualitative research studies about high stakes testing were analyzed by Au (2007). The author determined that in over 80% of the studies, the primary effect of using MCQ tests exclusively was a perceived narrowing of the curriculum to fit the test. Participants also changed their teaching practices to a more teacher-centered approach to cover all of the needed content.

The use of MCQ Tests as an Assessment Method in Nursing Education

Teacher developed MCQ tests are the primary method of student assessment in undergraduate nursing education. In many programs, it is the only method of assessment (Walloch, 2006; Oermann, Saewert, Charasika, & Yarbrough, 2009). To understand why teacher-developed MCQ tests are preferred and so highly valued, it is important to understand the history of nursing licensure.

The National Council of State Boards of Nursing's (NCSBN) NCLEX-RN was created in 1978 and is considered the impetus for using MCQ examinations as the primary method of student assessment (Billings & Halstead, 2009). In 1994, the NCSBN launched the first fully computer adapted licensing examination (Benefiel, 2011). One of the test designers explained that this type of testing allowed for increased security, and more importantly, the ability to ascertain entry level competency based upon the difficulty of the questions, not just the amount of correct answers (Wendt, 1999).

The first question on the examination is different for all applicants and is set at the passing standard; the next question depends upon how the applicant answers that first question. For example, if the first question is answered incorrectly, the computer will

give the tester a question at a lower cognitive level. Then, as the tester correctly answers additional questions the computer will gradually return to questions that reflect the passing standard. There is no mechanism for a candidate to skip or revisit previous questions, which is how the computer adapts to the tester. All passing questions are written at the applying and analyzing levels. To pass this test, candidates must answer 50% of these types of questions correctly. There is a maximum of 265 questions presented over a 6 hour time frame. Fifteen questions are not counted but are being evaluated for future examinations. It is possible to pass this test with as few as 75 questions. The NCLEX-RN was designed by nursing content and testing experts to meet strict psychometric standards (NCLEX Psychometrics, 2011). Continuous research is conducted to establish cognitive levels, readability, content validity, consistency, objectivity, and reliability of the examination and of every MCQ. The NCLEX-RN is one of the most rigorously researched high stakes standardized examination and has a pool of over 3,000 questions (NCLEX Psychometrics, 2011; Wendt & Harmes, 2009; Wendt & Kenny, 2009; Wendt & Kenny, 2007; Wo & Gorham, 2010).

Well-constructed teacher-developed MCQ examinations can assess higher cognitive levels and critical thinking. Morrison and Free (2001) described best practices for writing MCQ that promote and test critical thinking. The authors included 50 highly discriminating stems that faculty could adapt to test at the application and analysis levels and assess clinical judgment. Their work has been used to design a standardized exit examination which is considered by some researchers highly predictive of NCLEX-RN success (Lavandera et. al, 2011; Nibert, Young, & Adamson, 2008).

In numerous case studies (Su & Osisek, 2011; Su, Osisek, Montgomery, & Pellar, 2009; Su, 2007; Su, Osisek, & Starnes, 2004), Su and associates described item writing and test construction practices at a large Bachelor of Nursing program. The authors noted that writing context dependent items that tested critical thinking was difficult, and explained that nursing faculty are often not well prepared for this process. The authors outlined their experiences and best practices for revising nursing examinations. As a result of their practices, the quality of test items as measured by item difficulty and item discrimination values improved.

The critical thinking skills of first semester nursing students were also enhanced following a test taking skills workshop and exposure to MCQ tests written at higher cognitive levels (Schroeder, 2007). In a quasi-experimental research study using a pre-test, post-test design, all nursing students enrolled in an associate degree program completed a standardized critical thinking examination at the beginning of the first semester nursing course. Later in the semester, a test taking workshop was offered by the researcher. Following the semester examinations, both groups completed a second standardized critical thinking test. The researcher hypothesized that students in the experimental group who participated in both a test taking workshop and had experiences with teacher-developed tests at higher cognitive levels would demonstrate improved critical thinking. The author noted that while the critical thinking scores improved on the second standardized critical thinking examination for students who participated in the test taking workshop and completed the course examinations; there was no corresponding improvement in scores on the teacher-developed MCQ examinations.

While the NCLEX-RN and other standardized examinations are reliable and valid, teacher-developed tests have historically been poorly designed (Notar, Zuelke, Wilson, & Yunker, 2004). In an early study to determine the cognitive levels of MCQs on teacher-developed tests in nursing programs, Cross (2000) analyzed 110 final exams from randomly selected nursing schools in the United States. The author examined 130 exams from 66 programs and determined 92% of the questions submitted were multiple-choice in design. Three experienced nurse educators determined the cognitive levels and found that inter-rater agreement was negatively correlated, which supported research that nursing faculty have difficulty correctly identifying the cognitive levels of MCQ. Over half (51.3%) of the questions purporting to test at higher levels of Bloom's taxonomy by the submitting programs were actually from the knowledge domain, and more importantly, only 6.2% were at the analysis level.

In an unpublished study, De Pew (2001) researched how often nursing faculty determined the reliability and validity of MCQ examinations and whether a relationship existed between this practice and ultimate student success on the NCLEX-RN. Over 1,000 nursing faculty from 326 nursing programs were surveyed, and De Pew found no relationship between reliability and validity determinations and ultimate student success. In a graduate thesis which used a descriptive survey design and a researcher-developed tool, Rowland (2005) asked faculty in associate degree programs to identify their item writing and test construction practices. The results were surprising: only 43.7% of respondents consistently used a test blueprint, and only 41.5% consistently categorized questions according to Bloom's taxonomy (L. Rowland, personal communication, May

11, 2011). Faculty expressed the belief that MCQ examinations were the best method to prepare students for the NCLEX-RN and cited as the preferred student assessment by undergraduate nursing faculty in a descriptive survey research doctoral study conducted by Walloch (2006).

In an unpublished doctoral dissertation using a descriptive correlational design, Killingsworth (2013) found that nursing faculty do not routinely use test construction best practices when writing test items or designing examinations. The author surveyed faculty from Bachelor's in Nursing program located in 31 states and discovered that many do not blueprint examinations based upon the NCLEX test plan, peer review test items, conduct extensive item analyses, assess linguistic bias of questions or use current evidence based nursing practice when writing test items.

Nurse educators used a quality improvement method to evaluate faculty designed MCQ examinations. Clifton and Schriner (2010) analyzed the final examinations from three nursing courses in a baccalaureate nursing program. The number of questions totaled 256 and 20 questions from each examination were randomly selected for analysis. While purporting to test students at the applying and analyzing levels, 54% of the sample questions were determined to be from the remembering and understanding cognitive levels. Additionally, the difficulty values of 63% of questions were determined to be "too easy" (p. 14) and 30% of distracters were not plausible and therefore not selected by students.

The issue of implausible or non-functioning distracters and the impact upon the reliability and validity of the test has been studied extensively by Rodriquez (2005). In a

lengthy meta-analysis research studies, he determined that three distractors were as reliable and valid as the traditional four distractor model. This research has been confirmed by numerous studies in the nursing education literature (Redmond, Hartigan-Rogers, & Cobbert, 2012; Tarrant, Knierim, Hayes, & Ware, 2006; Tarrant & Ware, 2008, 2010; Tarrant, Ware, & Mohammed, 2009). Using both Rodriquez (2005) and Tarrant and Ware's (2010) conceptual framework, Redmond et al (2012) found no statistical difference between using three or four distractors in a teacher-developed multiple choice examination.

Masters et al. (2001) conducted the only review of nursing test banks in the educational literature and analyzed the MCQs accompanying nursing textbooks for adherence to Haladya and Dowling's (1989) multiple-choice question guidelines and Bloom's taxonomy (1956). Nearly 3,000 questions were assessed and the authors found over 2,000 violations of best practices. While most of these were minor, the researchers found that only 6.5% of all of the questions were written at the analysis level while 46% were written at the knowledge level, a non-passing standard on the NCLEX-RN.

While researching how undergraduate nursing students used textbooks, Costanzo (2009) learned that the textbook was used primarily by students to prepare for MCQ examinations. Using focus groups, Costanzo ascertained that students often used the textbooks to discover what types of questions faculty may ask on tests or to challenge faculty during test reviews. Students expressed frustration at the "tricky" and "confusing" wording of MCQ examinations and concerns that the answers on the teacher-developed tests often contradicted the textbook. Students questioned whether scores on MCQ tests

accurately proved nursing competence and were upset that teachers relied so heavily on test-banks.

The Impact of Flawed MCQ Examinations

Other authors have described how poorly designed MCQ examinations penalize students. Dowling (2005) analyzed four basic science examinations in a medical college and found that flawed MCQ comprised between 36 to 65% of questions analyzed, and these questions ultimately disadvantaged medical students. Dowling concluded "some students-perhaps as high as 10-15%...were incorrectly classified as failed when they should have been classified as passed, due solely to flawed item formats and the ineptitude of test item writers" (p. 141). This outcome reinforces the need for educators to create test blueprints and to conduct item analyses to determine content validity and reliability of teacher-developed examinations.

Tarrant, Knierim, Hayes, and Ware (2006) described similar issues after an extensive analysis of nursing MCQ examinations. Over a five year period, 2,770 questions were analyzed and 46% violated item writing guidelines, and over 90% tested at lower cognitive levels. In a later study (2008) Tarrant and Ware found that issues with MCQ item writing penalized the higher achieving students more than borderline students.

More recently, researchers (DiBattista & Kurzawa, 2011) analyzed 1198 MCQ items from 16 classroom tests from 12 different courses in an undergraduate university. Nearly 50% of all courses in this college used MCQ examinations for student grading at least one third to one half of the time. In their findings, the larger the class size the more often MCQ assessments were employed. The authors noted multiple issues with test item

construction and statistical analysis of the items. One third of the reviewed items had poorly functioning distractors and 30% did not meet published discrimination index standards. The authors noted that if poorly performing items were removed from testing the scores of 96% of the test takers would increase

Flawed MCQ items can even affect the quality of assessment tests for professional development. Stagnaro-Green and Dowling (2006) reviewed 40 MCQ tests included with continuing medical education articles in the *New England Journal of Medicine*. The researchers used Haladya and Dowling's taxonomy (2002) and determined that nearly every question examined had three item-writing flaws which could potentially impact the reliability and validity of the tests.

In a replication pilot study of Tarrant and Ware's (2006) research, Nedeau-Cayo and partners (2013) reviewed and analyzed over 2,000 MCQ questions designed by nursing professional development specialists for staff nurse education. The researchers determined that 34.6% of questions contained over two item writing flaws and 90% were written at Bloom's knowledge and comprehension levels. Over 85% of the questions contained at least one common item writing flaw. The authors noted that a grade of 100% was required to attain competency on each test therefore the staff had "developed work around methods" to attain the needed grade and based upon this study questioned whether staff had learned the content or were just "test savvy" (Nedeau-Cayo et al., 2013, p. 56).

The Use of MCQ Examinations with Students who Have Limited English Proficiency

Student issues with reading and comprehending the English language can impact the understanding of MCQ examinations. In a metasynthesis of ten qualitative research articles into the problems of nursing students with limited English proficiency, Starr (2009) explained that these students may have command of conversational English, but during school are suddenly required to learn two new languages, Standard English and health-related terminology. This impacts students' ability to read, comprehend, and analyze teacher-developed MCQ questions at the higher levels of Bloom's taxonomy. Many foreign educated students were prepared in schools that valued memorization over conceptualization, making test questions that test critical thinking and application more difficult.

This contention was supported by an extensive review of all literature related to the educational needs of Asian nursing students in American nursing programs. Using the search parameters of "retention strategies for ESL nursing students", "teaching strategies for culturally diverse nursing students" and "barriers of the ESL nursing student", Scheele, Pruitt, Johnson, and Xu (2011) searched seven databases and found 55 articles addressing the needs of culturally diverse students. Their results confirmed that students who have limited English proficiency often struggle during examinations and first time NCLEX-RN pass rates were significantly lower than for students with English as their primary language. The researchers concluded that current nursing educational practices were often inadequate to meet the needs of these students.

One of the earliest nursing educators to research this topic was Gardner (2005). In this phenomenological study, Gardner explored the experiences of racial and ethnic minorities in predominately white nursing programs. The 15 participants represented nine ethnic cultures and were between the ages of 22 and 47. Two of the students were men. Following data analysis, one theme that emerged was overcoming obstacles. Students said that while they were fluent in conversational English, reading highly technical language and writing papers were difficult. The participants also identified MCQ tests as a major obstacle, and cited feeling exhausted and overwhelmed during testing.

Using case studies, Lujan (2007) and Brown (2008) outlined faculty strategies to prevent linguistic bias in MCQ examinations for Mexican-American and foreign educated nursing students respectively, and suggested methods faculty could use to assist students with limited English proficiency become more successful test takers. Bosher (2006) analyzed 19 MCQ examinations from six nursing courses for a total of 673 test questions. The following criteria were assessed: (a) test-wise flaws, (b) flaws of irrelevant difficulty, (c) linguistic/structural bias; and (d) cultural bias. The results were significant, "an average of 2.2 flaws per item [and]…of the 52 types of flaws, 28 types occurred at least 10 times" (p. 263). In a follow up mixed-methods study, Bosher and Bowles (2008) found that modified test questions were rated as "more comprehensible" by 84% of students with limited English proficiency. The authors suggested 15 recommendations to make MCQ examinations fairer to all students.

In a qualitative study to determine the perceptions of ethnically diverse nurses about barriers to completing their education, Amaro, Abriam-Yago and Yoder (2006) interviewed 17 participants (14 women, 3 men). The sample included 8 Asian nurses, 4 Latino nurses, 2 Portuguese nurses, 2 African-American nurses, and 1 African nurse. Thematic analysis of the interviews determined non-native English speakers often need to translate written content to their native language before they can understand it. The participants felt that using this practice penalized them and made taking MCQ tests more difficult.

Foreign-born baccalaureate nursing students in Texas were asked about stress and perceived faculty support in an interpretive phenomenological study using a mixed methods approach. One of the themes which emerged was language issues; all students experienced problems with reading comprehension and the need to translate English to their native language, which they noted impacted test taking ability (Junious, Malecha, Tart, & Young, 2010).

The issues surrounding MCQ testing and students with limited English proficiency are not limited to nursing education. Malau-Aduli (2011) explored the experiences and coping mechanisms of international medical students in Australia using both interviews and quantitative review of examinations. Test analyses determined significant variations in student performances on MCQ tests. During interviews, the participants confirmed that language issues were often difficult to overcome and impacted both scientific writing and test taking.

Not only are students impacted by this issue. Donnelly, Mc Kiel, and Hwang (2009) explored the perceptions of faculty teaching limited proficiency English nursing students. The authors wanted to determine what the faculty believed were the factors that influenced the success or failure of these students. Qualitative thematic analysis supported that non-native speakers of English had difficulty understanding test questions or completing course work due strictly to language issues. Faculty often felt frustrated and unable to determine the best way to assist their students with limited English proficiency.

Summary of Review of Literature

Multiple-choice question examinations have been the primary method to assess students' knowledge of theory in nursing programs with mixed results. Research clearly demonstrates that well-constructed MCQs can assess the higher levels of cognitive knowledge and be an excellent assessment of student learning. However, problems with objectivity, readability, validity, and reliability can arise with poorly constructed teacher-developed tests. Determining if students are meeting learning outcomes may be difficult if only one method of assessment is used. Nurses need to demonstrate clinical judgment, be problem-solvers, and communicate effectively in today's rapidly changing health care environment (IOM, 2010; QSEN, 2013). The research is inconclusive as to whether faculty are able to adequately assess these characteristics using only teacher-developed MCQ examinations.

Research has confirmed that teacher-developed MCQ examinations may not be fair to all learners. Nursing experts are calling for a more diverse nursing force; yet, non-

traditional students and those with limited English proficiency often have difficulty with MCQ assessments, which can impact progression in nursing programs. Nursing faculty are quite frankly in a quandary, and question if using additional methods to assess learning will ultimately affect initial NCLEX-RN pass rates. It is unclear whether innovative assessment methods that reflect student learning styles and newer pedagogies can coexist with the need to prepare students for high-stakes licensure examinations that use MCQ written at the applying and analyzing levels. This problem was the focus of a heated debate at a nursing conference I attended (NLN Summit, September 21, 2011) as faculty from across the United States discussed the need to educate a more culturally diverse and non-traditional student population while still retaining accreditation, which is so closely tied to initial NCLEX-RN pass rates.

Implications

Understanding the teacher and student issues related to using only teacher-developed MCQ examinations will help the faculty to ultimately determine the best teaching and learning practices for this program. If the faculty are committed to this method of student assessment, then professional development in the area of item writing and test construction is imperative. Faculty will also need to be mentored and coached by more seasoned faculty who have expertise in MCQ test development.

Research into this area may also demonstrate a need for changes in the program's curriculum and overall assessment practices. For example, if faculty expresses the need to use other assessment methods in addition to teacher-developed MCQ examinations, then education in the areas of evidence-based assessment is needed. This could be

addressed with workshops or seminars, or peer mentoring or cognitive coaching (Maskey, 2009). Content about concept maps, reflective writing, portfolios, peer evaluation, clinical simulation, rubrics, and other innovative assessment methods would need to be included.

Summary

Despite research supporting the use of multiple methods to assess student learning (Chappius, Stiggins, & Chappius, 2012), nursing programs continue to predominantly rely upon teacher-developed MCQ examinations to evaluate students' knowledge of nursing theory and to assign grades. While today's nursing students are more diverse and nontraditional, assessment practices in nursing programs have changed little in the past 20 years. Recently, nursing education experts have called for a radical change in student assessment methods which would reflect newer and more innovative pedagogies and student learning styles.

The problem is not with all MCQ examinations as research has clearly demonstrated that standardized high stakes tests such as the NCLEX-RN are objective and reliable. The issue is with the use of teacher-developed tests that do not meet the same strict standards. This practice can potentially penalize students and prevent progression in nursing programs.

Faculty in one associate degree nursing program have determined that the best way to evaluate students in the theory portion of each course is teacher-developed MCQ examinations. This has resulted in excellent NCLEX-RN initial pass rates, but it may also be a possible reason for a high student attrition rate. Many diverse and non-

traditional students who excel in the clinical component of each course have failed the program due to the inability to pass MCQ tests.

In Section 2, I have outlined the research findings of this doctoral project study. Case studies are considered the most appropriate design for doctoral studies that research or explore problems on a local level. A case study is used when the phenomenon in question is difficult to research in isolation since it is so closely tied to many other variables (Yin, 2009). Faculty assessment practices are linked to teaching methods, student learning outcomes, student retention and attrition, the curriculum design, and total program outcomes, therefore isolating one or two variables would have been difficult. Section 2 also includes data collection criteria and tools, participants, data analysis methods, a discussion about the protection of research participants and methods to ensure the reliability and validity of data analysis, as well as the results of categorical analysis of the data.

I wanted to develop a thick, rich description of this issue and therefore used multiple sources of evidence including interviews and focus groups with key participants, document review, and journaling. During the data collection period, course coordinators involved with test construction were asked to maintain a diary describing their feelings prior to, during, and following each examination and outlining any issues with each test. The test blueprints, cognitive levels of randomly selected questions, and the item analysis of all MCQ tests administered during the data collection period were also reviewed.

In Section 3, I describe my proposed project. I addressed the faculty's perceived issues with teacher-developed tests as the only method to evaluate students in the theory

component of each nursing course. Data analysis determined the best plan to address these issues was faculty development in the areas of student assessment and evaluation. A three day workshop will give both novice and seasoned faculty an opportunity to practice MCQ item writing, construct tests, and learn newer, more innovative assessment practices. To ensure transfer of learning, four monthly lunch and learns are planned as an easy and efficient way to reinforce this content and give the faculty the opportunity to discuss ongoing issues or concerns in a comfortable, nonthreatening setting.

Section 4 contains my reflections upon my proposed project. It includes my assessment of the strengths and weakness of the proposed project study, and implications for further research. It also explains how this capstone project impacted me as an educational researcher, practitioner, scholar, and project developer.

The faculty of this nursing program want to use evidence-based teaching and learning methods, yet were fearful that changing the current practice will cause a decrease in first time NCLEX-RN pass rate, a critical component for accreditation. The purpose of this research was to explore faculty and student issues related to using MCQ examinations as the only method to assess and grade students in the theory component of each nursing course.

Section 2: The Methodology

Introduction

The purpose of this research was to describe the issues related to using teacher-developed MCQ examinations as the only method for student assessment and evaluation in the theory component of nursing courses at an associate degree nursing program. Methodologists assert that case studies illuminate phenomena (e.g., Yin, 2009) or tell a unique story, and were categorized by Stake (1978, 2005) as intrinsic, instrumental, or collective. Intrinsic case studies focus on one unique phenomenon (Stake, 2005); in this case, the student assessment practices at this nursing program.

In this study, I used an intrinsic exploratory descriptive case study because this method describes or attempts to understand a problem. Merriam (2009) noted that a descriptive case study design allows the researcher to provide an in-depth and "thick description" (p. 43) to explain the interplay of all of variables within a real world setting. Little was known about how the teacher-developed MCQ examinations shaped the practice of the faculty, affected the students, or impacted the nursing program as a whole. This research was designed as a first step to understand these issues and provide a thorough contextual description of this phenomenon.

There are numerous case study designs described in the literature; however, the model conceptualized by Yin (2009) was the framework that guided this study. Yin stressed that the selection of a research design depends upon three criteria: (a) the question, (b) the research focus, and (c) the extent of researcher control. Case studies are most appropriate when the researcher is posing questions that "ask how or why… [or] the

investigator has little control over the events, and the focus is on a contemporary phenomenon within a real-life context" (Yin, 2009, p. 2). While this research question asked *what*, Yin theorized that *what* questions are also appropriate for a case study because what asks an exploratory question within the boundaries of "unit of analysis" (p. 29).

The focus of this research was a real situation bounded by time and place which met the second and third of Yin's (2009) criteria. The unit of analysis, or case in question, was the assessment practices of the nursing faculty at this associate degree program from the time it was established to the fall 2012 semester. There were few variables that could be controlled or manipulated in this setting as assessment practices are enmeshed within all aspects of the curriculum and educational program.

A descriptive survey research design was originally considered for this study. Asking the nursing faculty to complete a survey would have been a quick and efficient method to describe the issues involved. Using surveys would have other advantages as well. They can be administered anonymously by written, verbal, or electronic methods, are relatively easy to analyze, and if appropriately designed are highly valid and reliable (Creswell, 2008). However, surveys are also inflexible and can rarely explain or capture contextual issues (Colorado State University, 2011). As Creswell cautioned "survey data is self-reported information, reporting only what people think rather than what they do" (p. 414). Therefore, using a descriptive survey design would have limited the information obtained and failed to give a full, rich description of the problem.

There were advantages and disadvantages to using a case study approach. Case studies occur at a site illustrative of a problem and therefore provide vivid examples within the context and environment of that problem. Case studies afford the analysis of both qualitative and quantitative data from many sources. Using a case study design gave me the opportunity to provide rich descriptions and explore the problem over time. This was also one of its major drawbacks. Case studies are often quite lengthy and prolonged; and Creswell (2008) noted that single case studies lacked rigor and generalizability. Another problem was that participants may not be open and honest during interviews or have forgotten vital information (Creswell, 2008; Yin, 2009).

While other methodologists debate the rigor of case study designs, Yin (2009) considered case study design as one of the most difficult research methodologies to implement. This is because it requires the researcher to be flexible, adaptive, unbiased, and have good question writing and listening skills.

Participants

The case in question was the assessment practices of nursing faculty at an associate degree nursing program; therefore, the faculty and administration of the program were the unit of analysis and served as the participants. Yin (2009) called this a single case design, and described rationales for using this method. Rationales include the uniqueness of the case, the ability of the case to "capture the circumstances or conditions of an everyday or commonplace situation" (p. 48), or the ability of the case to reveal new information or insights. Because little is known about the practice of using only one

method to assess students and to assign grades in nursing programs, a single case study method was the best method to explore these phenomena.

I purposively selected participants from the faculty and administration of the nursing program. These individuals constructed the MCQ tests and made the decision to use those tests as the only method of student assessment. In addition, they were responsible for all policies and procedures and to systematically evaluate the curriculum of the nursing program. I invited all of the 13 faculty members who had had student assessment responsibilities since the opening of the nursing program to participate. I also asked the Director of Nursing Education and Program Coordinator, who were members of the program administration to participate.

Gaining Access and Establishing Researcher Participant Relationship

Because I am a member of this faculty, it was not difficult to gain access to this site. I knew the gatekeepers and had already discussed with them the possibility of researching this topic. The Director of Nursing Education gave provisional consent and following Institutional Review Board (IRB) approval from Walden University and the hospital employing the faculty, final consent was obtained and data collection began. A letter of agreement was developed and signed by the Director of Nursing Education which afforded me full access to needed documents and participants. I established the groundwork for this study, knew all the faculty participants, and understood the culture of the nursing program, students, and faculty.

Glesne (2011) called this backyard research and cautions that it can have both positive and negative ramifications for the researcher. I have been a faculty member at

this program since 2001, and have been asked by fellow faculty to mentor them in the area of MCQ item writing and test construction. I also served as the curriculum champion for the program's assessment and evaluation committee.

Being part of this bounded unit was potentially problematic as administration and faculty see me as part of the group and may have been confused by my researcher role. Knowing all of the participants and essentially being part of the problem had ethical implications as well. It can be difficult to remain objective and impartial, or to ask probing questions of coworkers.

While bracketing (Powers & Knapp, 2010) is considered a part of phenomenological research, I believed it was an essential aspect of this study to ensure fairness and objectivity on my part as the sole researcher. I needed to become aware of my own feelings, and recognize any biases about this topic since I am a member of this faculty, and actively involved in this issue. While I have my own strong core beliefs about this topic, I fully appreciated that I am a part of this unit of analysis and therefore needed to remain open minded about how my opinions may have colored the data collection or impacted the data analysis.

Ethical Protection of Participants

Because I am part of this faculty, there was the distinct possibility that participants may have perceived this research study as threatening or punitive; therefore, I needed to ensure that the rights of all faculty were protected. Prior to beginning this research, I submitted my proposal to the IRBs of both Walden University and the hospital which employs the faculty and administration of this program. All participants were

provided with information about the purpose and aims of this study and given the opportunity to ask me questions prior to consenting to participate. All participants signed informed consents (see Appendix C). The advantages and disadvantages to participating were clearly delineated on these consent forms. Participants were assured that their identities would be protected and masked, and that they had the right to leave the study at any time without consequences. After signing the Data Collection Agreement (see Appendix B), the Director of Nursing Education allowed me access to the password protected anonymously completed online student course end survey. This was the first step in data collection and it started during the fall 2012 semester.

Data Collection

Before data collection began, a case study protocol was developed and used during the data collection process. This protocol was more than just the questions to be asked. Yin (2009) stressed that it is the procedure that is used to guide the collection of data and will ultimately increase the reliability of the research. The protocol (see Appendix D) outlined the data collection plan and also included a full description of all participants to be interviewed, as well as documents to be reviewed and analyzed.

The principles of data collection were summarized by Yin (2009), and maintaining these principles helped me to reduce bias and ensure the reliability and validity of the case study research design. The first principle stressed is the "use of multiple sources of evidence" (Yin, 2009, p. 114). Using multiple sources of evidence allowed me to explore all of the factors and variables involved to begin to make meaning

of this phenomenon. Methods to collect this data included interviews, focus groups, personal diaries, and document review.

The next principle outlined by Yin (2009) was the creation of the database which described and explained how the data were organized. This involved creating a filing system and maintaining personal notations about the data in a format that could be easily retrieved. Notations were made on copies of interview or focus group transcripts and personal diaries, and then saved as Microsoft Word files or scanned into a computer file and saved electronically. Personal notations were also made on copies of the documents and any archival data which were reviewed and either saved as files or scanned into a computer and saved electronically. The creation of these databases was essential for determining the credibility and trustworthiness of the data and the establishment of the audit trail (Yin, 2009).

To increase the reliability of all case study information obtained, I saved and documented what Yin (2009) calls the "chain of evidence" (p. 122). I maintained case study notes, a journal of my own reflections of the data analysis, full transcripts of the interviews and focus groups, and copies of my review of the test blueprints, item analysis, and the cognitive levels of selected questions in a file on an external flash drive. These procedures will allow a reader to follow the data from initial collection to analyses and was important to ensure the trustworthiness of the research. Data will be kept in my home office in a locked file cabinet for a period of 5 years. After that time, all data will be deleted.

Interview

I conducted an individual interview with the Director of Nursing Education. According to Seidman (2006), the purpose of an interview is "an interest in understanding the lived experience of other people and the meaning they make of that experience" (p. 9). Assessing, examining, and understanding how the nursing education administrator viewed the current examinations and assessment methods was important to gain a full appreciation of this issue. The interview was semistructured in nature to allow me the opportunity to ask specific questions but still permitted the participant to fully describe, clarify, or explain her perceptions and beliefs (Merriam, 2009). The interview was 1 hour in length, scheduled after the review of documents, and conducted in the nursing administration offices. Interviewing the nursing program director, after reviewing any pertinent documents, allowed me to clarify any questions I may have had concerning information contained in the documents. The interview protocol is outlined in Appendix F. The interview was audiotaped and I took handwritten notes.

There were advantages and disadvantages to using an interview. Interviews can provide vivid and detailed information but are also limited by the participant's memories or shaded by her particular biases (Creswell, 2008). Using a semistructured format was potentially problematic as pertinent information could be missed if I neglected to ask good follow up questions. Yin (2009) cautioned that the researcher must be able to write insightful questions, follow the case study protocol, and listen intently all at the same time.

Focus Groups

Merriam (2009) defined a focus group as "an interview on a topic with a group of people who have knowledge of the topic" (p. 93). The focus groups were scheduled after the review of the documents, which allowed me to clarify information and ascertain the faculty's perceptions about the MCQ examinations they had developed. The focus groups were audiotaped using a digital recorder, were 1 hour in length, and conducted in the faculty conference room. I served as the moderator and took notes during the focus groups.

There were advantages to using focus groups. Focus groups were relatively easy to schedule, allowed me a degree of control over questioning, and afforded the participants a chance to provide in-depth descriptions of the issues (Creswell, 2008). The focus group protocol is contained in Appendix G. Patton (2002) and Merriam (2009) posited that focus groups are an excellent method for the construction of new knowledge. The open and free discussion permitted participants to hear and consider other opinions, and ultimately allowed for reflection. This openness is also one of its major disadvantages as not all participants may be willing to be honest about the topic in question or to actively participate. Conducting focus groups can also be difficult and require a great deal of concentration and flexibility. As moderator, I needed to ask good questions, be patient, actively listen, and simultaneously record, observe, and make notations (Cresswell, 2008; Glesne, 2011).

Personal Diaries

During the data collection period two course coordinators were actively involved with test construction. I asked them to maintain a diary describing their feelings prior to, during, and following each examination (see Appendix E). Additionally, they were requested to make notations about any problems encountered during test design or administration, or issues with the test blueprints, item analysis, or student success. Yin (2009) noted that personal documents such as diaries are an excellent way to "corroborate and augment evidence" (p. 102) found by other data collection methods. Creswell (2009) noted that personal documents can be essential for understanding the phenomenon from the unique perspectives of the participants.

One of the selected course coordinators is responsible for the first semester fundamentals of nursing course (Nursing 1). Historically, there were few student comments about the quality of the teacher-made examinations in this course and student attrition remained at 4% from 2009 to 2011. However, in 2012, the attrition rate for this course jumped to 13% (Curriculum Committee Minutes, 2009-2012). Therefore, the insights of this faculty member about test design and construction while it was occurring were imperative to gain a full understanding of this issue.

The other course coordinator is responsible for the third semester course which has recently undergone a great deal of change. Nursing 3 is essentially three nursing courses in one as it focuses on obstetrics, pediatrics, and behavioral health nursing concepts. This course was unchanged from its initial class in 2007 until 2011. In this course students learn all of the concepts from the three nursing specialties

simultaneously. For example, during the first two weeks, content was "front loaded" so that all students have content related to all three areas to allow them to be effective in the clinical setting. In 2011, based upon faculty and student feedback, the course was restructured and redesigned. The students were divided into sections and over a five week time frame they received theory and content about only one of the nursing specialties. This resulted in three mini five week nursing courses, two 30 question tests in each section, and necessitated a different schedule for students and faculty from previous years.

Following the course, the student evaluations of the course structure and schedule were positive; however, many of the faculty believed their workload was impossible and felt test security was an issue after a few test questions were posted on a social networking site. Due to the fact that there were three sections, there needed to be three different versions of all tests. This caused the faculty to search for a pool of reliable and valid questions. For example, in the behavioral health portion, two faculty members designed all of the tests and they noted in a curriculum meeting that they "spent hours trying to find enough questions to prevent one group from giving the content to the next group" (Curriculum committee minutes, Spring 2012).

There were also issues related to the typing and administration of one of three versions of the final examination during the previous semester. During the actual examination it was noted by faculty that some distractors were bolded; therefore, some students had the correct answer for five questions. It was also noted that all of the answers for two medication dosage calculation questions were incorrect due to typographical errors. Based upon the workload issue, faculty voted to return to the original format for the fall of 2012. Due to these issues, the opinion and perceptions of this course coordinator about item writing and test construction were important to fully appreciate these problems.

Document Review

I reviewed test blueprints, item analyses, and the cognitive levels of a sample of faculty developed MCQ exams. Additionally, nursing students are asked each semester to complete faculty and course evaluations at the completion of the semester. These Likert-type surveys are completed confidentially on *Survey Monkey*® and yield both nominal and qualitative data. There are two statements which address the issue of student assessment and students often write long comments about MCQ examinations. Faculty names had been removed by the Program Coordinator; therefore, I was blinded to this information. All faculty were given this blinded information at Curriculum Meetings and it is used for the program's ongoing Systematic Plan of Evaluation (SPE Committee Minutes, 2012).

The Case Study Database and Categorical Analysis

The literature describes numerous approaches to case study data analysis (Merriam, 2009; Yin, 2009). Case study data analysis requires ongoing reflection upon the data and constant focus on "the theoretical propositions, developing case descriptions, using both quantitative and qualitative data and examining rival explanation" (Yin, 2009; p. 126). Merriam (2009) stressed that the biases and world-view of the researcher need to be "identified and monitored" (p. 15) as these shape the overall analysis of the data. In qualitative research, the researcher acts as the instrument and all data analysis is filtered through the researcher's theoretical framework or philosophy; therefore, ensuring the accuracy and credibility of the data analysis is imperative.

While there are no hard and fixed rules for case study data analysis and the development of the case study database, as the first step I organized all interviews, document reviews, and observations. I then categorized and removed all redundancies before intensive analysis began. This step involved data transcription, data coding, personal notations and reflections about data, identification of recurring patterns and themes, and developing and validating generalizations (Yin, 2009).

I found this initial step daunting. It required multiple steps to organize artifacts, documents and interviews. Yin (2009) suggested "playing" (p.129) with the data until the researcher finds a system that works. While qualitative computer assistive tools are available, Yin cautioned that these are often difficult to use with case studies as the data are more than just narrative in nature. I developed a matrix of categories that reflected my literature review. For example, students with limited English proficiency often

struggle with faculty developed multiple choice questions; therefore, any narratives or documents that reflected that concept were placed in that category. It was important to tie the data to the theoretical framework or literature review.

In case study research, the coding process is ongoing and concurrent with the data collection and is known as categorical analysis. Therefore, I organized the data by type and tentative categories initially. For example, I initially transcribed the interviews and began coding based upon the research question, subquestions, or the review of literature. I then coded the data from document analysis in the same manner, which allowed me to gradually explore and gain a general sense of the data. Merriam (2009) referred to this process as "open coding" (p, 178). The goal at this point was to read the interview and focus groups transcripts or documents thoroughly and assign possible names or codes for all of the data. This process yielded numerous codes so I needed to streamline them and began to describe and develop themes, which were ultimately layered and interrelated (Creswell, 2008).

For the next step, I developed the case description. This required me to write a full and complete description of the characteristics of this case and all of the data. Using multiple sources for data collection afforded me the ability to triangulate the data. Yin (2009) referred to this as "the development of converging lines of inquiry" (p. 115) and this approach assured the accuracy and credibility of the finding. Also, using multiple sources of evidence added to construct validity as each data source provided a measurement of the same phenomenon.

The coding process and data analysis included all narratives and descriptions of faculty perceptions. Discrepant or deviant cases are defined as information or themes that seem to contradict or oppose the initial hypothesis, or not support the overall analysis of data (Merriam, 2009). Since this was an intrinsic exploratory descriptive case study, all discrepant data were included and explained as this would ultimately give a clearer picture of all of the perceived faculty opinions and beliefs.

Evidence of Quality

Trustworthiness and Credibility

Qualitative methodologists (Merriam, 2009; Yin, 2009) described techniques to validate or establish the trustworthiness of the research findings. These included credibility, transferability, dependability, and confirmability (Merriam, 2009). In qualitative designs, the researcher is not interested in external validity or generalizability (Yin, 2009) but in transferability, or a determination that research findings are applicable for similar persons in similar situations.

The term trustworthiness is used in qualitative research when describing the accuracy and credibility of the research findings. Yin (2009) opined that the trustworthiness of the case is confirmed if the researcher used multiple sources of evidence, created a case study database, and maintained the chain of evidence. Those criteria were outlined above. In addition, I provided rich, thick descriptions of this case, and utilized triangulation and member checking to establish reliability and validity.

Rich, Thick Descriptions

Providing a rich, thick description of the case study's setting and participants lends credibility to the researcher's findings. Qualitative research reports should contain detailed descriptions of the case in question and development of themes with lengthy, rich narratives, and direct quotations from participants. Creswell (2008) suggested "writing strategies" (p. 193) which include the use of long or short quotations explaining the themes, actual conversations or full text of the interviews, intertwining researcher's reflections and observations, and tying the narratives to the literature or conceptual framework. In this study, a descriptive narrative with quotes from the interviews, focus groups, and document analysis were included to give the reader a contextual understanding of this problem.

Triangulation

Triangulation is the process of using multiple evidence sources and then comparing, contrasting, and corroborating the data from all of these sources (Yin, 2009). According to Yin (2009) comparing data from multiple sources helps to ensure "construct validity and reliability" (p. 114) as the developed themes need to exist in these multiple data sources (Creswell, 2009; Merriam, 2009). Yin noted that using triangulation adds to the credibility and dependability of a case study. The major drawbacks to using triangulation are the additional time required to fully compare and contrast large amounts of data and the possibility of bias on the part of the researcher (Thurmond, 2001).

Member Checking

Member checking entails asking participants to check the accuracy of the findings and to confirm if I clearly understood their perspectives. Following transcription of the interviews and focus groups I asked participants to review the transcript to confirm the accuracy of their comments. Member checking allowed me to make sure I fully understood participant comments and perceptions, and afforded additional clarification or reflection on the part of the participants. However, there is often a concern that "member checking relies on the assumption that there is a fixed truth of reality that can be accounted for by a researcher and confirmed by a respondent" (Robert Wood Johnson Foundation, 2008, para. 3). In the narrative, I included any discrepancies I found during member checks as I believed it would allow for a fuller, more complete description of these issues and would ultimately add to the credibility and dependability of this case study.

Assumptions, Limitation, and Scope of the Study and Delimitations

Assumptions

1. Participants would be open and honest during interviews and focus groups.
2. Participants' beliefs and perceptions would be representative of the nursing faculty as a whole.

Limitations

Single case study designs are potentially limited by concerns of trustworthiness (Yin, 2009). In this case, the research was conducted at one nursing program over the course of one semester and analyzed by one person. Therefore, there are risks to

credibility, dependability, transferability, and confirmability. Methods to address these issues were outlined in the previous section.

Scope and Delimitations

The scope or boundaries of this single case study were the assessment practices of faculty at this one associate degree nursing program since its inception in 2007. There was no attempt to study or describe teaching methods, teaching styles, learning activities, the overall curriculum, course objectives, or program outcomes.

Data Collection Procedure

Document Review

As the first step in data analysis, I reviewed all of the student comments from anonymously completed course-end evaluations. While the program coordinator had already summarized the student comments there had been no previous attempt to tie or link these comments to the overall assessment practices at this nursing program. A total of 12 course end surveys were available on *Survey Monkey* (Spring 2010-Fall 2012) and I made a decision to use only those surveys as they were typed and anonymously completed; therefore, I would not be distracted by handwritten student comments or be tempted to identify which student wrote the comments. Table 2 outlines information from these surveys.

Table 2

Student Course-end Comments

Year	Total number of comments	Total number about MCQ examinations
Spring 2010	86	31
Fall 2010	36	5
Spring 2011	71	29
Fall 2011	77	21
Spring 2012	67	18
Fall 2012	69	30
Total	406	134 (33%)

Next, I assessed the teacher-developed MCQ test blueprints, and reviewed actual test questions for conformity to established nursing education item writing standards (McDonald, 2013) and examined the item analyses. Since 2009, 50 examinations containing over 2000 MCQs were designed and administered to students. Traditionally, these examinations do not change significantly from semester to semester. Faculty usually only revise problem or dated questions, or change non-functioning distractors.

I made the decision to randomly select questions from this pool of over 2000 possible questions. Using a random generator (*Random.org*), I selected 10 numbers from 1-100 to accommodate 50 and 100 question examinations and 3 numbers from 1-33 to accommodate 33 question examinations which were instituted in 2011 in one course. For example, in spring 2010, the second semester course administered 4 examinations of 50 questions each and 1 final examination with 100 questions for a total of 300 questions.

Each faculty member who taught content for an examination contributed between 10 and 25 questions per test. Therefore, for a 50 question examination I randomly selected questions 2, 12, 29, 33, and 50 to insure that questions from all faculty members were included in the analysis. For a 33 question examination, questions 4, 11 and 25 were analyzed. A few of the final examinations contained 75 or 100 questions; I therefore used questions 2, 12, 29, 33, 50, 64 and 72. Questions on each test are grouped according to content. For example, in Nursing 2, the first exam contained 25 questions from respiratory content and 25 questions from cardiac content. Therefore, the random questions selected were representative of all theory content.

The third step in the document review process was an analysis of faculty maintained diaries. Two course coordinators maintained a diary during the data collection period and noted any issues with item writing, test construction, test administration, grading or student challenges.

Interview

I conducted an individual interview with the Director of Nursing and the protocol is contained in Appendix F. The Program Coordinator was asked to participate but was not available during the data collection period. The interview was semistructured in nature which allowed me the opportunity to ask specific questions but still permitted the director to fully describe, clarify, or explain her perceptions and beliefs (Merriam, 2009).

I designed the interview with the Director of Nursing to elicit broad comments about the overall assessment of student learning at the nursing program. The questions focused on assessment practices, formative versus summative assessment, attrition, and

the faculty's comfort using alternative assessment methods. The interview was one hour in length, scheduled after the review of documents, before the faculty focus groups, and conducted in the nursing administration offices. This allowed me to clarify any questions I had concerning information gleaned from the other data collection methods. The interview was audiotaped and I also took handwritten notes. The interview protocol is outlined in Appendix F, and following the interview, a verbatim transcript was shared with the Director. She made no corrections or clarifications of the content.

Faculty Focus Groups

Nine of the 13 full time faculty members who are responsible for test construction agreed to participate. Due to scheduling conflicts, two focus groups were needed. Six faculty participated in the first focus group and the remaining three participated in the second group two days later. For the purpose of writing this narrative, faculty members have been given the following pseudonyms: Beth, Coleen, Danielle, Diane, Jane, Laura, Megan, Sharon, and Sandra. Faculty demographic data is included in The Case Study Protocol (see Appendix D) and Table 3.

Table 3

Participants

Name	Position	Credentials	Years as nurse educator
Nadine	Director of Nursing Education	MSN RN Certified Nurse Educator (CNE), EdD credits (*ABD)	27
Danielle	Course Coordinator	MSN, BSN, CNE	20
Diane	Course Coordinator	MSN, BSN, Geriatric Clinical	22

		Specialist	
Sharon	Course Coordinator	MSN, BSN, CNE	11
Coleen	Course Coordinator	MSN, BSN	35
Megan	Faculty Member	MSN, BSN	4
Sandra	Faculty Member	MSN, BSN	8
		Pediatric Nurse Practitioner	
Beth	Faculty Member	MSN, BSN	8
Jane	Faculty Member	MSN, BSN	20
Laura	Faculty Member	MSN, BSN	1

Note. * = all but dissertation.

During the focus groups with faculty I asked a few of the same questions I asked the director. The remaining questions were narrower in scope and were designed to gather information about how the faculty perceived item writing and test construction. Faculty were also asked questions about whether they believed that using only one method of student assessment impacted the teaching/learning process at the nursing program. The questions were also intended to understand how faculty viewed the process of formative and summative assessment.

I scheduled the focus groups after review of the documents which allowed me to clarify information and ascertain the faculty's perceptions about the MCQ examinations they had developed. The focus groups were audiotaped using a digital recorder, were one hour in length, and conducted in the faculty conference room. Verbatim transcripts of the focus groups were shared with participants; there were no corrections or clarifications submitted.

Data Analysis Results

The overriding research question for this exploratory case study was: What are the issues related to the use of teacher-developed MCQ examinations as the only method of

student assessment and evaluation in the theory component of courses at a suburban associate degree nursing program? As previously noted, data was obtained from a review of student course-end evaluations, a sampling of actual examination questions, an interview, two focus groups and faculty maintained diaries.

Throughout the data analysis process, I recorded all data as per the case study protocol outlined in Appendix D. As required by Yin (2009), I maintained this database and my ongoing categorical analyses to ensure the trustworthiness of the research and to establish the audit trail.

Summary of Review of Couse-End Evaluations and Exam Questions

Student Course-End Evaluations

I collected a total of 406 comments from 12 student course-end surveys, and of these 134 (33%) comments mentioned assessment practices or MCQ examinations. These comments were then coded and initially placed into eight categories. After extensive reflection and "playing with the data" (Yin, 2009, p. 129), I categorized student issues with the current assessment practices into three broad categories:

- Disconnect between course content and questions,
- Disconnect between study guides and questions,
- Quality of the MCQ examinations.

Disconnect between content and questions. There were 47 student comments related to this category. All comments were lengthy and directed at individual faculty members. Only five comments were positive and expressed appreciation for faculty questions that were reflective of content that was fully discussed and stressed during

lecture or other classroom activities. Other students perceived a "disconnect" between what was stressed in class or clinical experiences and what was actually tested on unit examinations. They also noted that while "critical thinking" was stressed by all faculty; many questions required memorization of facts and not application or analysis of nursing concepts. One comment in particular summarized this category: "Her test questions, for my opinion, don't help us much with critical thinking, which from what I learned is the most important. Most are knowledge level questions…I am confused about memorization versus critical thinking."

Disconnect between study guides and exams. Students often request study guides prior to MCQ examinations to help them focus and prepare for the actual examination. There are no policies or procedures at the nursing program regarding this practice and individual faculty members decide whether or not to provide this information. A few faculty members post sample MCQs, some may give topic overviews (blueprint), while others give an outline for students to complete as they study. For many students, these study guides become the focus of all examination preparation.

There were 20 student comments on the 10 course-end evaluations about study guides. There were only three positive comments thanking individual faculty members for the format and design of these guides. The remaining 17 comments were negative and some were quite personal in nature. Student comments focused on the need to study massive amounts of material yet they believed they were tested on content that was not stressed in class or contained on the "study guide." One student wrote the following which was reflective of all of the negative comments, "Study guides could be more

related to the exams. At times, I felt the study guides were guiding me towards the wrong direction. I feel the study guides only helped me with 1 or 2 questions per test."

Quality of MCQ exams. Based upon previous review of the course-end evaluations following each semester, I had formulated the opinion that the majority of student comments focused on editorial issues and the overall quality of the MCQ. However, after this in-depth analysis, there were only 24 comments that mentioned this issue. Additionally, there were no comments at all about inability to read or understand the faculty developed MCQ from students for whom English is a second or additive language. This was a surprising finding since nursing research has focused extensively on this issue (Bosher, 2008; Gardner, 2005) and this is an area of concern for the program's faculty. The majority of comments noted the following problems on faculty developed examinations:

- Typographical errors and other editorial issues
- Confusing or ambiguous stems or distractors
- Incorrect key distractors based upon textbook or faculty lecture

Two student comments summarized their perspective about the design of MCQs:

> "I'm not saying we shouldn't have to take tests but, some of the questions were horrendous. I know that the point of some questions is to pick the best choice out of 4 right answers. But, there were several questions that had wording that we weren't taught (that the instructor would know based on her job experience) and I don't know how I would have any knowledge of those terms when I learned it the way it was taught and I have never had a nursing job. The

> teachers don't feel the need to admit they're wrong so the students who studied 10 hours for an exam don't get the points because the teacher wrote a bad question and doesn't want to admit it."
>
> "I feel that more time needs to be spent writing the test questions. I feel that [an instructor's] questions had nothing to do with the material taught and that there was frequently more than one BEST answer for the question. The wording on the tests should be more student friendly."

Analysis of Examination Questions

As previously described, I randomly selected questions from all available MCQ examinations from the fall semester of 2009 until the fall semester of 2012 for a total of 50 tests and analyzed them for conformity to established item writing standards. Using the criteria outlined by McDonald (2013), Clifton and Schriner (2010), and Tarrant and Ware (2012), I reviewed the cognitive level, the item analysis, the item difficulty, and the relationship to course objectives of the randomly selected questions.

When I decided to do a random selection of questions, I was unaware that many of the questions were the same and fell in the same order on each test from year to year. This meant that for some examinations, the same questions were analyzed two or three times. For example, in the first semester course there is a stable bank of test questions that are used from year to year with little change. The course coordinator requests that faculty not totally change questions, but to rework stems and change poorly written or non-functioning distractors as needed. Therefore question 12 on Test 1 from 2009 was essentially the same as question 12 on Test 1 from 2011. However, in other courses,

questions are reordered or removed; therefore, in those instances, I reviewed different questions on every test for each year.

After reviewing a total of 250 questions from all four semesters, I noted the following practices which did not conform to established MCQ item writing or test analyses standards:

1. Inconsistent test blueprinting,
2. Disconnect between stated cognitive level and actual cognitive level of questions,
3. Nonfunctioning distractors, unclear key, or multiple possible correct answers,
4. Lack of clinical vignettes,
5. Editorial issues,
6. Use of poorly written questions taken directly from textbook test banks or other materials,
7. Use of poorly performing questions from year to year,
8. Inconsistent examination item analysis (Haladyna, Downing, & Rodriquez, 2002; Mc Donald, 2013; Tarrant & Ware, 2012).

Test blueprinting. None of the nursing course coordinators consistently documented a test blueprint or required individual faculty members to submit the test blueprint to them for review prior to each examination. A total blueprint of all tests with course specific objectives was also unavailable for three of the four nursing courses. When test blueprints were requested by me for review, only one course coordinator was able to give me a complete blueprint for all written test questions. However, course

coordinators did have information about the MCQ cognitive level and relationship to the NCLEX-RN test plan for most of the examinations.

MCQ cognitive levels. As with the test blueprints, not every question analyzed had a cognitive level listed, and for some of the questions, the cognitive level did not conform to the actual cognitive level of the question according to Bloom's revised taxonomy (Anderson & Krathwohl, 2001). Many of the questions reviewed were incorrectly identified as testing at the applying or analyzing cognitive levels were actually testing at the recall or understanding level. This practice has been clearly noted in the nursing education research (Cross, 2000; De Pew, 2001; Rowland, 2005).

Nonfunctioning or implausible distractors. According to Haladyna, Downing, and Rodriguez (2002), there is no agreement in the literature about the best or correct number of distractors in a MCQ. In published guidelines they noted that "effective choices" (p. 312) should be developed and that research has supported the use of three well written, highly discriminating selections instead of the traditional four distractor choices (Tarrant & Ware, 2012). In this case, in 20% of the 250 randomly selected questions, two or more of the four distractors were not selected by students.

Lack of clinical vignettes. Nursing education item writing experts stress the importance of presenting clinical vignettes in every question which require students to demonstrate clinical judgment and critical thinking. The stem should include a patient situation, a nurse, and data that are pertinent to answer the question (Tarrent & Ware, 2012). Course coordinators stressed the importance of having a nurse-patient interaction

in the stem of each question, and while this practice was improving, four of the questions analyzed did not conform to this practice.

Poorly written questions. Included in this category are questions that did not conform to established item writing standards for the following reasons:

1. Not all distractors were the same length or the key was longer than the other options,
2. Vague or confusing questions or questions that are not relevant to nursing practice,
3. Extraneous information in the stem or "window dressing,"
4. Clues in the stem or distractors,
5. Negative wording (Haladyna & Downing, 2002; Tarrent & Ware, 2012).

Summary of the Individual Interview, Focus Groups, and Diaries

Interview

During the interview, the Director confirmed what has been noted in the nursing education literature when she expressed her concerns that poorly written or designed tests may have impacted student attrition. She noted that questions on some examinations were "unclear" or "difficult for students" to read or comprehend. She wondered if students who were academically borderline were needlessly penalized by poorly designed examinations or by using only one assessment method.

The Director also stressed the idea that students needed to know the rationales for all questions that were answered incorrectly on examinations and supported the practice of reviewing all tests with students. She strongly believed that this would assist students

to understand needed content and that it was ultimately the faculty's responsibility to explain in depth why an answer was correct or incorrect. She worried that some test reviews had become confrontational and emphasized that the reviews needed to be conducted in a positive and professional manner.

It was interesting to hear the Director state that she "struggled" with item writing and test construction when she was a faculty member. She noted that she always asked colleagues to review her questions and was a strong proponent of peer review and collaboration during item writing and test construction. She stated she believed that the entire process of the assessment of student learning was better than in the past but there was still "room for improvement."

Focus Groups

The faculty participants discussed many issues during the focus groups. The faculty appeared genuinely concerned about best practices and expressed frustration about time constraints and other obstacles which often made item writing and test construction difficult. Both experienced and inexperienced faculty made comments that reinforced the students' perceptions concerning issues with the quality of some of the MCQs used. Most of the faculty did not create original MCQ test items but relied instead on test banks, published NCLEX-RN review books, and other resources for ideas, stems, and distractors. They stated they "rewrote" or "revised" or "tweaked" the questions until they were happy that the item included needed information to assess the students' understanding of content.

Faculty stated that they believed that many students were only focused on individual exams and there was little retention of content from test to test or from semester to semester. Faculty noted that the high stakes nature of course examinations may contribute to this student practice. Coleen noted "I know we focus on test-taking strategies, I have to focus on helping students understand how to take a test because that is what their grade is based on." Jane shared "I think it does affect how we teach…I know during a test and I see one of my questions I wonder, did I get the main points across, did I stress that content, was it covered well in class?" Others noted that they tried not to "teach to the test" but it was difficult not to be influenced by the knowledge that certain content was indeed tested and will ultimately determine if a student will progress in the nursing program.

Some faculty believed the practice of changing questions from year to year based upon one year's item analysis affected the overall reliability and validity of all tests. They also discussed methods they had used in the past to improve the overall quality of examinations. The faculty assigned to the third semester course noted that they had instituted a practice of peer reviewing all test questions. While most members felt this practice had helped the overall quality of the examinations, two members felt that all this "tweaking and editing" actually caused more problems as important information was inadvertently deleted or changed in some questions.

Participants in both focus groups expressed concern about the entire test construction process as they thought it was time consuming and often frustrating. There were numerous comments about the overall quality of nursing test questions available in

test banks and the difficulty revising them to meet test construction standards. Many noted that finding the perfect "fourth distractor" was difficult. Faculty in both focus groups strongly believed that poorly designed questions impacted students negatively. The faculty also made numerous comments about how the quality of test questions affected them personally. They described feeling frustrated during test reviews when they realized that students did not understand a question or answered a questions incorrectly based upon a poorly designed stem or an inaccurate distractor. They did not like "throwing out" a question but sometimes they felt they had no choice to be fair to students.

Faculty also thought that using only exams to assess student learning and assign grades may impact the way they teach. During the discussion, the faculty noted that they needed to prepare students to take a high-stakes standardized examination which was primarily MCQs for licensure. Therefore, the NCLEX test plan and question style became the focus of their test construction. In the past, they used group projects, written assignments, and clinical performance to determine theory grades but believed these methods could be "unfair" or "subjective" and therefore made the decision to use only "objective" tests to assign course grades. Many noted that term papers or group assignments did not help students prepare for a high stakes licensing examination and that students needed practice taking similar format exams over the course of their nursing education.

Soliciting the faculty's opinions and views about the test reviews during the focus groups elicited many positive and negative comments. The entire faculty noted that they

believed that reviewing tests and explaining rationales for correct and answers was an important part of the learning process. However, the current method impacted their relationship with students as the reviews often became confrontational as students "fought" for every point. They observed that this situation was due to the high stakes nature of each examination. One comment from Sandra illustrated the faculty's opinion of test reviews. She noted "we have debates and fights for grades, the students think that if they complain enough the grade will be changed." Sharon noted that she is committed to giving students feedback but "some students have an answer in their heads and no matter what I say they do not see that the answer they selected is incorrect."

Faculty also noted that working with students individually revealed information about how students studied, processed information, and ultimately prepared for course examinations. During individual examination reviews faculty discovered that students were trying to memorize information and many had no idea how to study effectively. Students would arrive at the individual review sessions with entire chapters "highlighted" or with the entire Power Point class presentation memorized. These students could explain facts or understand data; however, they often could not apply these facts in a clinical situation. Three comments were especially revealing about student study and examination preparation habits.

Diane noted,

> "What struggling students need is a one on one session not a large group review. They need to understand how they arrived at an answer. What was their thought process? How they are reading the question? If a student gets seven out

of ten questions wrong and they realize that they were making up stories or picking the second best answer, doesn't that help them more than just knowing rationales?"

Sharon shared another perspective. She commented that students often "don't think about how A+B=C; they memorize A, and B, and C, but may not conceptualize how it is all related…maybe we set them up for this by the way we test, I don't know." Danielle noted another aspect of student preparation. She told the focus group a story about a student. "When I sit down to do a one to one review it is so interesting. This is when we can really see how they [students] think and how they read and answer questions." She went on to describe a student who works as a pharmacy tech who had failed a test. She noted, "He got three pharmacology questions wrong and after only a few minutes I realize that he was thinking like a tech, not like a nurse."

The focus group participants were also asked if using one method to assess student learning and assign grades impacted how students studied. The faculty surmised that students often attempted to memorize content rather than conceptualize or critically think. They noted that there was little retention of content from one test to the next and an inability to "apply theory in the clinical setting." The faculty also expressed the idea that students needed to accept responsibility for their own study habits. The majority of faculty stated that testing became all about blaming the faculty and not an opportunity for feedback or an assessment of learning.

All of the participants explained that while they may struggle with test construction they believed that the entire process was improving and that as a whole the

current tests are much better than when the program was started. Some faculty commented that the current assessment method may not be working for all of the students based upon the student attrition rate from the nursing program (see Table 1); yet noted that despite this high attrition rate, the initial NCLEX-RN pass rate was higher than the average for other associate degree nursing schools throughout the country which reflected the overall quality of the nursing program.

The majority of faculty expressed support for the current method of assessing student learning and assigning grades. Five of the faculty believed perceived that the current MCQ exams were formative in nature. One faculty member noted that she believed that all of the tests were formative as faculty were preparing students to take a standardized examination to attain licensure. However, the remaining faculty members and the Director of Nursing Education noted that there was no opportunity for formative assessment that was not graded due to the structure of the courses and time constraints.

Faculty noted that they attempted formative assessment in the past by using classroom assessments methods such as an audience response system, One Minute Papers and Muddiest Points (Angelo & Cross, 1993), but some noted this process was not useful. Faculty also had concerns about integrating other methods to assess learning or assign grades. The majority of faculty believed that since the NCLEX-RN examination was "an objective examination", assessment methods within the nursing program should be similar.

A few faculty did not feel comfortable grading written assignments and stressed that it was often difficult to remain objective or impartial. Some were unfamiliar with

using rubrics to grade written assignments and others believed that written assignments ultimately did not help students pass the licensure examination. All four nursing courses use a final summative examination; therefore, the majority of the faculty perceived each test as a formative appraisal of student learning.

Faculty Diaries

As I analyzed the diaries maintained by two faculty members (Sharon and Danielle), I found clear examples of issues with the quality of the faculty-developed examinations. Sharon outlined student and faculty issues related to test construction as she reflected upon the examinations from her course. One issue that surfaced was the amount of student transcription errors as answers were copied onto the scoring sheets. She noted that students often changed answers from correct to incorrect or transcribed the wrong answer on the tally sheet. According to her journal, this practice occurred on every test. In one incident, a student made a transcription error which resulted in 10 points being deducted from the grade.

Sharon also noted that on every test there were questions that were poorly worded or contained confusing distractors which caused some questions to be discarded or required that multiple answers be accepted. She tracked a total of 5 examinations which were administered to a class of 79 students. The following issues were noted on each examination:

- Errors on the computerized scoring key,
- Errors of transcription on the computerized scoring key

- Unclear or poorly written stems or distractors which required accepting two answers or the elimination of the question.

In her diary, this course coordinator noted that when she saw some of the examination results, she "felt sick to my stomach; I had never seen this type of result in my 10 years as course coordinator." She also noted that she did not feel comfortable requiring other faculty who taught with her to change questions and noted "if there are problems with their questions; ultimately they have to defend them to students."

Danielle, the course coordinator for the first semester course maintained a diary as well. She shared that maintaining the diary caused her to analyze the test results more closely and to perform an in-depth assessment of how students answered each and every question. For the first time since becoming an educator, she compared every student's computer answer sheet to answers on the actual test and decided that she would continue this practice in the future.

Traditionally, this course has had few issues with item writing and test construction and this was also true this semester. Unlike the examinations in the other course, there were only a few questions that needed to have two answers accepted or eliminated. The first semester students seemed to have difficulty with "select all the apply questions." This type of multiple response question is now common on the NCLEX-RN licensing examination therefore, the faculty include a few on each examination.

Categories Identified from Interview, Focus Groups, and Diaries

Following the analysis of the interview, focus groups, and submitted diaries, I identified the following categories that may have impacted the assessment of student learning:

- Faculty workload and time constraints,
- Few opportunities for formative assessment,
- Student focus on memorization not conceptualization of content,
- Insufficient availability of well-constructed MCQs in test banks or other resources,
- Inadequate understanding of item writing and test construction best practices by some faculty,
- Focus on initial NCLEX-RN examination initial pass rate.

Patterns, Relationships, and Themes

After lengthy reflection upon the categorical analysis and a thorough review of my own journals and code book, I was able to ultimately develop patterns, relationships, and themes. As previously noted, analysis of the student-course end surveys, faculty constructed examinations, an interview, faculty focus groups, and faculty maintained diaries yielded the following categories:

- Disconnect between course content and examination questions,
- Disconnect between study guides and examination questions,
- Inconsistent examination blueprinting,
- Incorrect identified MCQ cognitive level,

- Questions containing non-functioning distractors or lacking clinical vignettes,
- Poorly written questions that do not conform to established standards,
- Faculty workload and time constraints,
- Few opportunities for formative assessment,
- Student focus on memorization not conceptualization of content,
- Insufficient availability of well-constructed MCQs in test banks or other resources,
- Inadequate understanding of item writing and test construction best practices by some faculty,
- Focus on initial NCLEX-RN examination initial pass rate.

As required by Yin's (2009) exploratory case study methodology, I analyzed the above categories for the existence and identification of patterns. This step was followed by linking the patterns to form relationships and determining themes from these relationships. Finally, the themes were aligned with the research question. Throughout this process, I continually looked for data that were different or discrepant. All of the data supported that there were issues with the current assessment practices at the nursing program.

Patterns

When I reviewed all data sources, I found similar issues with the current state of student assessment at the nursing program. The students identified perceived "disconnects" between course content, study guides, and examinations, and made over 60

comments to that effect. While the students described these discrepancies as a "disconnect," the faculty's perspective was one of students' inability to study effectively and students' reliance on "memorization" not "conceptualization." Student comments demonstrated a reliance on faculty "study guides" to prepare for nursing examinations. In over 20 comments, students identified problems with the design and quality of examination items. An analysis of a sample of test questions, conversations with faculty, and a review of faculty maintained diaries supported this. Students also requested "additional methods" for grading and faculty noted that formative assessment and/or additional methods to assess student learning were difficult to include in the format of the current curriculum. The faculty noted their focus was to prepare students to ultimately pass a high-stakes standardized licensing examination and believed that the use of faculty-developed primarily MCQ examinations supported this outcome.

Relationships

During data analysis, I found an emphasis on the relationship between the use of only faculty developed examinations to assign grades and the inability to formatively assess student learning. Faculty noted that due to the structure and format of each nursing course it was nearly impossible to "revisit previously learned" content or provide adequate feedback to students. Students felt that this lack of feedback and issues with item writing and test-construction contributed to the high-stakes nature of each nursing examination. Students' inability to answer questions constructed at the applying and analyzing cognitive level was often related to ineffective student preparation and study habits. Students had difficulty answering questions that required exploration of clinical

data or the application of nursing theory in a clinical situation. These issues impacted the teaching/learning process as many students were focused entirely on the course examinations; not on the NCLEX, the ultimate standardized examination that would allow them to practice as registered nurses. The high-stakes nature of each examination affected the teacher/student relationship as the students were focused only on grades and often believed that the exams were unfair. Exam days were stressful for both faculty and students and exam reviews occasionally became confrontational.

Themes

After reflection upon the data, categorical analysis, review of my own journals and code book, and examination of the two theoretical frameworks guiding this research, the following themes emerged which described the issues related to the use of teacher-developed MCQ examinations as the only method of student assessment and evaluation. The exclusive use of faculty developed MCQ examinations to assess student learning and assign grades has limited the:

- Opportunity for formative assessment of student learning,
- Faculty's ability to assess higher order thinking; and,
- Development of metacognition of the part of students.

Limited opportunities for formative assessment. In section one of this paper, formative assessment was defined as "activities used by the teacher to determine a student's level of knowledge and understanding for the purpose of providing students with feedback and planning future instruction. The feedback and future instruction may be concerned with remediation or the provision of further learning opportunities"

(Baroudi, 2007; p. 39). Both the faculty and students' comments described few opportunities between the faculty developed MCQ examinations to provide clarification of course content or feedback to students. Each week, new nursing topics are introduced and examinations occur every four to five weeks depending upon the course. Some faculty have tried to use quick classroom assessment methods but that practice was not standardized or always successful.

Educational experts have called for the use of evidence-based formative assessments to determine student learning (Knight, 2006; Nicol & Macfarlane-Dick, 2006; Stiggins & Du Four, 2009; Clarke & Dede, 2010). Two major parts of this process are feedback and reflection. This feedback is supposed to be designed to assist students in assimilating content before being tested. However, often the only feedback students receive about their understanding of nursing theory is during the faculty conducted posttest reviews. Assessment experts (Mc Manus, 2008; Williams, 2011) debate whether a graded examination should be considered a formative assessment.

Limited faculty's ability to assess higher order thinking. The ability to construct or create new knowledge is the highest level of Bloom's revised taxonomy (Anderson & Krathwohl, 2001). In this level, students must draw upon previously learned knowledge or skills to create a new product, and it is associated with other concepts including generating, planning, and producing. Using only faculty developed MCQ examinations limits faculty in determining if students can critically think and demonstrate clinical judgment. Experts agree that it is impossible to use MCQ examinations or other selected response methods to assess or evaluate student's

performance in creating new knowledge (Chappuis, Stiggins, Chappuis, & Arter, 2012; Suskie, 2009). Assessment methods associated with this concept are assignments that require students to describe alternative hypotheses, outline the steps to solve a problem, or create a new product based upon specific guidelines and instructions (Suskie, 2009). Nurses must consider alternatives, solve complex clinical problems, and be creative in many patient care situations. As a prominent nurse educator posted on a blog, " I've always thought that to continue to use multiple choice [tests]…does a disservice to our students in promoting limited ways of learning and limited ways of demonstrating learning" (B. Thompson, September 27, 2006). A teacher-developed multiple choice question exam alone does not indicate which nursing students are critical thinkers and use clinical judgment in practice (Del Bueno, 2005).

Limited the development of metacognition on the part of students. One of the conceptual frameworks I used to guide the data analysis is Bloom's Revised Taxonomy (Anderson & Krathwohld, 2001). The authors' noted that multiple-choice question examinations can test at the application and analysis levels and some authors posit well-constructed questions will also test at the evaluation level (Oermann & Gaberson, 2009). However, no MCQ examination can test, assess, or measure metacognitive knowledge or students creating new knowledge (Anderson & Krathwohld, 2001).

Students wrote comments regarding a "disconnect" between what they perceived as important to study and the actual examinations. By relying on "study-guides" or memorization of faculty PowerPoint lecture slides, students did not adequately reflect upon the content, assimilate salient points, or transfer the information into the clinical

realm. To be successful in clinical situations, nurses need to think deliberately and critically and these are skills that cannot be adequately assessed by MCQ only. Metacognitive knowledge is student awareness and acceptance of responsibility for his or her own knowledge and thoughts; and has been simply defined as "thinking about thinking" (Walker, 2012). Martinez (2008) includes critical thinking as an aspect of metacognition. An educator needs to first understand how students think and reason before implementing strategies to promote clinical reasoning and clinical judgment. Nurses use metacognitive strategies in all clinical settings when they assess patient data, design appropriate interventions, and then critically reflect and evaluate if interventions were successful (Pointexter, Hagler, & Lindell, 2014).

Metacognitive knowledge includes three aspects (a) strategic knowledge, (b) cognitive knowledge, and (c) self-knowledge (Anderson & Krathwohl, 2001). During the faculty focus groups and in faculty diaries, the faculty described student scenarios of these aspects. Strategic knowledge is defined as "general strategies for learning, thinking, and problem solving" (Anderson & Krathwohl, 2001, p. 56). Diane alluded to this concept when she discussed her opinion of test reviews. She noted that it is more important for students to understand how they arrived at an answer and an appreciation for the thought process that impacted their conclusions. She believed that just a review of course content or providing rationales to questions did not help students transfer theory into the clinical setting.

The second aspect of metacognitive knowledge is an appreciation for context and conditions inherent in attaining knowledge. Students tend to think linearly and focus on

memorization to attain knowledge rather than conceptualizing what is learned. To attain cognitive knowledge, students need flexibility and adaptability in their thinking as situations or conditions change.

The third aspect of metacognitive knowledge is self-knowledge, which includes a student's awareness of personal thinking and studying strategies (Anderson & Krathwohl, 2001). Clapper (2011) explained for leaning to occur a student "must be able to reflect upon what they currently know and consider how new information is applicable to them" (Metacognition, p. 1). Faculty need to use methods to assist students to develop self-knowledge by providing honest and concrete feedback to assist students' learning. This idea was confirmed in the focus group data. One of the students interpreted and answered questions from a job-related perspective rather than demonstrating nursing judgment.

Ambrose and associates (2010) explained that strategies to promote metacognition are neglected in most educational programs and students often find themselves lost in higher education settings when they are given complex assignments or expected to apply or analyze options. The authors described metacognition as a cyclical process which ultimately allows students to "monitor and control their learning" (p. 192), and outlined five steps needed to become self-directed learners. They also delineated concrete methods for educators to develop metacognitive skills in their nursing students. Two of the strategies outlined are relevant for this nursing program. The authors encouraged faculty to provide students with performance based assessments early in the unit of study and with opportunities for self-assessment as these strategies increase

metacognition on the part of students. These are two areas identified as lacking in this nursing program.

The authors of The National League for Nursing's Core Competencies for Nurse Educators (2005) stressed the need for faculty to use varied, diverse, and evidence-based assessment methods to assess student learning in the cognitive, psychomotor, and affective domains. Nursing is a practice discipline; therefore, assessment methods that link theory to the clinical setting are needed. Using only one method to evaluate students and assign grades limits faculty in assessing students' critical thinking, clinical judgment, and professional competency. (Benner, Sutphen, Leonard, & Day, 2010; Utley, 2011). By using only one method to assess student learning and assign grades, opportunities for authentic assessment are missed. In a landmark nursing education study, Benner et al. (2011) determined that nursing programs need to vary their assessment methods to assess student learning and forgo the often singular emphasis on MCQ examinations.

Alignment of Themes with Research Question

The overriding research question for this exploratory case study was: What are the issues related to the use of teacher-developed MCQ examinations as the only method of student assessment and evaluation in the theory component of courses at a suburban associate degree nursing program? There were three subquestions:

1. How has using only teacher-developed MCQ examinations impacted the teaching/learning process?
2. How has using only teacher-developed MCQ examinations impacted the teacher/ student relationship?

3. How are students assessed formatively in the theory component of the nursing courses?

Impact on the teaching/learning process. Using only faculty-developed primarily MCQ question examinations to assess student learning and assign grades impacted the teaching learning process. One of the themes which emerged was that there were limited opportunities for faculty to assess higher order thinking or encourage metacognition on the part of the students. Both students and faculty agree that they are spending most of their time focused on either constructing or administering examinations. Since each examination is summative for an area of nursing content, the students perceive them as high-stakes. Students often direct their energy to memorization of facts and not conceptualization of salient points. This emphasis impacts learning as students are focused only on the unit examinations and often cannot remember content from test to test or semester to semester.

The faculty are also concerned at the amount of time and effort spent trying to "cover content" that will be tested or teach "test-taking skills" to assist students. They noted that many students are only focused on testing opportunities and have difficulty synthesizing nursing theory content. Many faculty want to offer additional assignments to assist students to attain higher grades, yet are concerned that these assignments are not objective and may result in lowering achievement standards. Faculty know that they are expected to maintain an excellent NCLEX-RN initial pass rate required for ongoing accreditation. Many faculty are not confident using alternative assessments they perceive such as subjective. All of the faculty noted that item writing and test construction

requires specialized skills to develop fair, unbiased, and reliable tests that will ultimately assist students to be successful on the licensure examination.

Impact on the teacher/student relationship. Due to the high-stakes nature of every nursing examination, the relationship between students and faculty is often strained during each testing opportunity. Examination administration and reviews often become contentious as students contest for every point and may not be receptive to faculty explanations about each question's rationale. Students often spend hours trying to find evidence in textbooks or classroom content to challenge the faculty's answers and rationales. When questions are not discarded, some students become angry and are not open to learning.

Formative assessment of student learning. My analysis of the evidence from all data sources supported few opportunities for formative assessment of learning and constructive feedback. Students did not perceive classroom assessment methods as helpful, and the faculty reported that there was little time for non-graded informal formative assessment in the theory portion of each nursing course. The faculty believed real formative assessment of student learning needed to be enhanced and developed.

Trustworthiness of Outcomes and Evidence of Quality

To ensure the trustworthiness of this research, I followed the criteria as previously outlined in the Evidence of Quality Section of this paper. I reviewed multiple sources of data (student course-end evaluations, actual faculty-developed examinations, diaries, and interview and focus groups transcripts) and all evidence supported that there were issues with the assessment of student learning practices in the program. Throughout the data

analysis process, I followed my protocol and created a complex database and case description. All documents, focus groups and interview narratives were analyzed, triangulated, and thick, rich descriptions were developed. An example of my codebook is found in Appendix H. All data, transcripts, and tape recording are on file in a locked cabinet in my home office.

Conclusion

To fully describe the issues surrounding the use of teacher-developed MCQ examinations as the only method of student assessment in the theory component of nursing courses, I completed an intrinsic case study. Case studies are an excellent method to provide a detailed, contextual narrative of a group or event and are often the first step in the process of understanding issues or changing practices within an educational unit.

Using a case study design was the best method for exploring problems and ongoing issues within a natural setting and when the researcher had little control over the variables. The assessment practices of this nursing program are linked to individual faculty beliefs, the mission and philosophy of the nursing program, and the educational objectives; therefore collecting data from multiple sources was important to fully explore this issue.

The participants for this case study included the faculty and the administration of the nursing program. I collected data by using one interview, two focus groups, two personal diaries, and a review of the faculty developed tests and student end of course surveys. Data were analyzed using categorical analysis.

Following categorical analysis and triangulation of the data obtained from the aforementioned methods, the following issues associated with using MCQ examinations as the only method of student assessment and evaluation in the theory component of courses at a suburban associate degree nursing program emerged. Using only faculty-developed primarily MCQ examinations:

- Limited the opportunities for formative assessment of student learning ,
- Limited the faculty's ability to assess higher order thinking; and,
- Limited the development of metacognition of the part of students.

During the focus groups, the faculty expressed the desire to include additional assessment methods to evaluate students and assign grades but were concerned that other methods would not be objective. While the majority of faculty expressed support for the current practice; they also realized that using only faculty-developed primarily MCQ examinations may not provide a fair assessment of student learning. Therefore the proposed project is a three day (21 hour) faculty workshop followed by four one hour monthly lunch and learns over the course of one semester which will be explained in depth in Section 3 and the full workshop content is included in Appendix A.

This workshop will serve two purposes. It will provide all faculty, seasoned and novice, the opportunity to review, critique, and ultimately revise the faculty-developed MCQ examinations to ensure conformity to item writing and test construction best practice. It will also introduce the faculty to authentic assessment and highlight other methods to assess and evaluate students and assign grades. The monthly lunch and learns

will help to promote transfer of learning and encourage the faculty to continue to grow and develop in this area.

Reading their diaries, listening to their voices, and reflecting upon the overall positive student course-end evaluations helped me to appreciate that the faculty of this nursing program are determined to provide a quality educational experience for all students. The program has an excellent reputation and the faculty are considered experts in the clinical setting. They have also demonstrated scholarship and professionalism by publishing, presenting at educational conferences, mentoring nursing education graduate students, and maintaining the accreditation status of the nursing program. Since 2009, the initial NCLEX-RN pass rate of the nursing program far exceeds the national average for associate degree nursing programs nationally. This case study was an opportunity to explore an issue that has vexed faculty and potentially impacted student attrition. The proposed professional development workshop will be the next step to assist faculty to continue to use evidence-based educational strategies and assessment methods.

Section 3: The Project

Introduction

The faculty and administration of this community college nursing program were committed to using evidence-based teaching and learning strategies and providing an excellent educational experience for students. To that end, they have maintained the accreditation status of the program within the state (Board of Nursing Accreditation Report, 2012) and nationally via the National League for Nursing Accreditation Commission (NLNAC, 2009). According to the former Director of Nursing Education at this program, accreditation reviewers commented that the primary reason for the effectiveness of the program was the clinical expertise of the faculty and their overwhelming commitment to student success.

Despite this commitment to using best practices, the faculty at this nursing program had been committed to using teacher-developed multiple choice examinations as the only method to assess student learning and assign grades. Following this intrinsic case study and categorical analysis I determined that this practice:

- Limited the opportunity for formative assessment of student learning,
- Limited the faculty's ability to assess higher order thinking; and,
- Limited the development of metacognition of the part of students.

To address these concerns, an overview of the faculty professional development (FPD) program about evidence-based assessment practices in higher education is presented in this section and Appendix A. The program consists of a 3 day workshop of six sessions (21 hours) followed by four, 1 hour lunch and learns; for a total of 25 hours.

During the faculty focus groups which occurred during the data collection phase of this project, numerous participants noted that they were uncomfortable using any assessment method that was not perceived as objective. They also stated that they needed some development in the areas of formative assessment and creating rubrics. The workshop was designed to meet these identified needs.

Goals of the Proposed Project

I designed the faculty development program to review multiple-choice examination best practices and introduce the faculty to authentic assessment methods. It will also contain an overview of formative assessment methods and allow faculty to design learning tasks and assignments that will facilitate this practice. It is the goal of this program to give both novice and experienced faculty enhanced knowledge, cognitive skills, and problem solving abilities as related to student assessment. It was designed to give faculty an appreciation of the ethical responsibilities and legal implications of student assessment methods.

The goals of the workshop are to:

- Review the best practices for item writing and test construction,
- Explain newer evidence-based methods to assess student learning and assign grades, and
- Describe strategies to assist faculty to enhance students' ability to use metacognition and reflection.

The program and participant learning outcomes and all aspects of the workshop are outlined in the Program Evaluation Section and Appendix A. There are nine program

learning outcomes and individual objectives created for each session. The literature delineates numerous models for designing educational workshops and seminars (Laureate Education, 2009). I used Vella's (2008) *Seven Design Steps* to provide the structure and format for this workshop because this model gave me the ability to create clear, succinct, and easily evaluated student learning outcomes.

Scholarly Rationale for Selection of Faculty Development

One of the conceptual frameworks guiding this capstone project is the National League for Nursing's Core Competencies of Nurse Educators (2005). Competency 3, which underscored the need for nursing faculty to use evidenced based assessment and evaluation strategies, was the framework for the research portion of this paper. Competency 6 stressed the need for academic faculty to pursue continuous quality improvement in the nurse educator role. This competency further stressed that "ongoing commitment to develop and maintain competence in the role is essential" (p. 6) and outlined eight tasks to assist faculty to attain expertise as nurse educators. These competencies were discussed in Section 1 of this paper and stress lifelong learning, professional development, and mentoring.

According to the NLN Position Statement (2002), many faculty have excellent clinical expertise but lack the needed skills to "facilitate learning, advance the total development and socialization of the learner, design appropriate learning experience, and evaluate learning outcomes" (p. 2). It also stated that while nurse educators need clinical expertise as nursing is a practice discipline, more attention should to be given to educator

role; therefore, all nurse educators need to be attentive to ongoing professional development in educational methods.

Researchers who have examined the assessment practices of teachers have confirmed that they often feel unprepared to adequately assess student learning outcomes (Frey & Schmitt, 2007, 2012; Mertler, 2009; Popham, 2009). This was evident at this nursing program. During the focus groups, the majority of the nursing faculty stated that they did not feel adequately prepared to use additional assessment methods to assess student learning and assign grades in the theory portion of nursing courses. They also expressed concerns that including additional assessment methods will allow borderline students to pass nursing courses while being unable to ultimately pass the NCLEX-RN examination.

As noted by the NLN (2001), "some individuals come to their faculty positions knowing little if anything about the principles of teaching and learning…A vast majority of nursing faculty have been prepared as advanced clinicians, not educators" (NLN Position statement, 2001, p. 2). In this case, nine out of the 13 full time faculty were not prepared as nurse educators at the graduate level but as advance practice nurses, clinical specialists, and administrators.

Based upon the culture of the nursing program and statements made during the focus groups, I believe this faculty will be open and receptive to learning from each other during a professional development workshop. Annually, the entire faculty attend at least two to three educationally focused professional development programs either on campus

through the college's Center for the Enrichment of Learning and Teaching (CELT), area hospitals, or through various professional nursing education organizations.

Additionally, the administration of the program is committed to faculty professional development. Money is budgeted each year for faculty to attend national, regional, and state level nursing and higher education conferences. There is also a vibrant and active Professional Development Committee which designs and conducts one program each semester for the faculty. For example, in January 2013, members of the faculty presented a program that included the following topics: the flipped classroom, simulation strategies, and active learning methods to promote student participation. Also, in May, 2013, the administration of the nursing program purchased a subscription to *Nurse Tim*, a web based nursing education professional development site which offers live and pre-recorded sessions on nursing education topics. A webinar from this site will be part of the workshop.

Members of the faculty have also presented nursing education or clinically focused sessions or posters at state and regional conferences with the financial support of administration. Two members of the faculty have written and published extensively and serve as reviewers for clinical and education focused peer reviewed nursing journals. Five members of the faculty and two administrators have attained the prestigious Certified Nurse Educator (CNE) title. It is clearly evident that life-long learning is promoted and is valued by this faculty.

Scholarly Rationale for Addressing the Problem

I wanted to design a professional development program that would assist the faculty attain needed expertise in the assessment of student learning outcomes and the design of MCQ examinations. All of the professional development programs attended by the faculty on this topic in the past were passive in nature. By this I mean participants essentially sat in a large convention center or room and listened to an expert in item writing and test construction describe best practices. Once the faculty returned to the nursing program there was no opportunity, incentive, or time allotted to develop or practice these newly acquired skills. The faculty's time was devoted to the same routines and while some did review or revise their test questions, others kept the status quo. This was primarily due to time constraints. This was clearly validated by the comments made during the faculty focus groups.

A three day professional development program using a workshop method followed by four monthly lunch and learns to ensure learning transfer fits the culture and climate of the nursing program. All of the faculty have been involved with peer mentoring and peer review. Caffarella (2002) defined a workshop as an "intensive group activity that emphasizes the development of individual skills and competencies in a defined content area" (p. 240). Workshop participants are typically experienced practitioners who are tasked to solve a problem in small groups (University of Kansas, 2013). Workshops evolved from a constructivist theoretical framework. Successful workshops focus on the contextual aspects of learning and are designed to explain new educational methods and allow participants to immediately practice new skills. The main

focus is on problem solving and the creation of new knowledge by the learner (Brandon & All, 2010; Maheshwari, 2012).

This workshop will be an occasion for faculty to work collaboratively in small groups reviewing and revising MCQ examinations in a relaxed non-threatening environment. It will afford them the opportunity to create content specific test blueprints that will guide the construction of all examinations. It will also provide an introduction to the concept of authentic assessment (Wiggins, 1989) and a review of methods to formatively assess student learning.

Review of the Literature

Since this professional development program is devoted to student assessment practices I needed to review multiple topics before designing the actual workshop. First, I reviewed all resources and papers from the Walden University EDUC 8104: *Facilitating Adult Learning Course*. Based upon the data analysis (Section 2), there are numerous issues related to using only one method to assess students and assign grades. Therefore, in addition to researching professional development literature, I needed to search for research about the concepts of authentic assessment, formative assessment, professional development, and best practices for multiple-choice item writing and test construction.

I conducted a literature search using the same method as my original review of the literature as outlined in Section 1. I requested articles from 2009 to 2013 from the following databases at Walden University Library: Academic Research Complete, Educational Resources Information Center (ERIC), Educational Research Complete,

Cumulative Index to Nursing and Allied Health Literature (CINAHL) and Medline Simultaneous, Walden Thoreau, and Google Scholar. The following keywords and combinations of terms were used: *faculty professional development, theories of faculty professional development, faculty professional development* and *nursing education, faculty professional development* and *higher education, formative assessment, authentic assessment, assessment of student learning, MCQ item writing and test construction best practices* and *educational workshops*.

Many of the classic articles about research into faculty professional development were written before 2009; therefore to use primary resources I needed to broaden the time line. Additionally, there were few relevant articles obtained about nursing faculty professional development as most focused on peer mentoring or the orientation of new clinical or adjunct faculty. Therefore, most of the reviewed literature is from the higher education setting. When the same articles or research were consistently cited in each database, I knew I had reached saturation.

Faculty Professional Development

Professional development is a broad and all-encompassing concept. It includes both formal and informal strategies designed to promote learning (Cox & Mayorga, 2010; Drumond-Young et al., 2010; Mundy, Kupczynski, Ellis, & Salgado, 2012; Shagrir, 2012). Methods of faculty professional development (FPD) include but are not limited to workshops, seminars, conferences, peer mentoring and coaching, role modeling, reflective practices, communities of learning, and membership in professional associations. Faculty professional development is defined as activities that are designed

to enhance the individual faculty member's role as a teacher, a scholar, a practitioner, or a person. Faculty development is usually designed to enhance the scholarship of teaching and learning (POD, 2007). The American Nurses Association supports this definition and noted that lifelong learning was necessary to attain professional goals and remain a competent educational and clinical practitioner (ANA, 2010).

Much of the research into FPD deals with faculty satisfaction with or the design of individual programs, rather than the impact these programs have on learning or student outcomes (Light, Calkins, Luna, & Drane, 2009; Webster-Wright, 2009). In an extensive review of the literature related to effective professional development and its impact on educational practices, Wayne, Yoon, Zhu, Cronen, and Garet (2008) argued that there is a dearth of randomized, control studies to guide best practices and stressed that research into effectiveness rather than simply faculty satisfaction was needed. This sentiment was echoed in an editorial (Draper & Clark, 2007) in a leading nursing education journal when they noted that most nursing professional development programs focused not on enhancement of student learning outcomes but faculty satisfaction.

Measuring improved student outcomes or changes in teaching practices is often difficult. In a uniquely designed research study, Ebert-May and associates (2011) found that following a professional development workshop about active learning strategies to teach scientific concepts, 89% of participants surveyed reported a change in teaching practices. However, when participants were videotaped in the classroom, only 25% had made an actual change in practice, the remaining 75% were still employing lecture and teacher-centered classroom methods.

Programs are more effective when teachers found that the content is coherent with their own learning needs and have the perception that newly learned strategies can be implemented immediately (Eib & Miller, 2006; Penuel, Fishman, Yamaguchi, & Gallagher, 2007). Shargir (2012) explained that while FPD can be viewed as taxing or arduous, participants in a mixed methods study perceived it as necessary to become proficient as teachers. A week long faculty workshop about technology integration and e-learning was evaluated for five years to determine effectiveness (Kenny, Banerjee, & Newcombe, 2010). The authors determined that a change in practices only occurred when faculty saw a direct link to ongoing student success. The authors noted that PD was more efficacious when adult learning theory guided implementation as faculty needed the opportunity to immediately implement and integrate the new learning in a supportive environment.

There are numerous models for FPD in the literature. One model was developed by O'Sullivan and Irby (2011) for academic medical education. Their review of the literature and best practices resulted in a new framework which they called a "faculty development community" (p. 424). The authors stressed that FPD does not occur in isolation and each faculty member is part of two distinct communities, the faculty community and the workplace community. The authors noted that FPD is only successful when there is synergism between these two communities.

Another model of faculty professional development was described by Mc Namara, Roat, and Kemper (2012). The authors' outlined an interdisciplinary approach that helped keep the focus on the learning needs of the nursing faculty. The nursing

program facilitators' hired a faculty professional developer who was responsible for the creation of a center dedicated to the educational needs both new and experienced full-time and adjunct faculty. The faculty have been committed to peer-mentoring, coaching and designed a classroom observation tool to guide assigned mentors and their mentees.

Best practices for FPD have been studied extensively by researchers at the University Of Pennsylvania Graduate School Of Education. In numerous articles (Birman, Desimone, Porter, & Garet, 2000; Desimone, 2009; Desimone, 2011; Desimone & Smith, 2011; Hochberg & Desimone, 2010; Rowley, Desimone, Porter, Garet, Yoon, & Birman, 2002) the authors have studied the role and impact of teacher professional development and based upon an extensive review of the literature described five concepts that are integral for effective professional development.

Desimone (2002, 2009, 2011) argued that the reason for FPD should not be just an increase in teacher job satisfaction and stressed that all professional development should stimulate a change in practice which will ultimately result in a change in student outcomes. This is an important aspect of the program I am planning. I will not consider it a success if the faculty simply state that they "enjoyed" the workshop or "learned new ideas;" the success of this workshop will ultimately be evaluated by a change in the practices at the nursing program and integration of newer assessment methods into the curriculum. Vella (2002) called this concept the impact of the program. The five core concepts of professional development which have been supported by the research are described below. After the identification of these core concepts, Desimore (2009)

continued research into professional development and constructed a conceptual framework. This framework guided the creation of the FDP workshop.

Content focused. Effective programs focus on subject matter and practices that describe how students best learn that content. In this planned workshop, the faculty of the nursing program will use relevant resources and reference materials to create new student assessments. They will be actively writing and revising questions, creating rubrics, and ultimately critiquing and evaluating each other's work. Therefore, the faculty will be immediately engaged and able to put the new learning into context and practice.

Active learning. Desimone's (2009, 2011) exhaustive review of the professional development literature has demonstrated that the most effective programs are those which involve the learner and allow them to be active participants. This has been the missing piece from all earlier professional development sessions attended by the faculty. Previously, faculty were passive participants and there was not an opportunity to immediately practice or implement the newly learned content. Transfer of learning will be facilitated by the four one hour lunch and learn sessions scheduled during the semester following the original workshop.

Coherence. This is an important aspect. The program must align with the learners' goals and be part of a complete set of ongoing learning opportunities. The program must be supported by both the faculty and the administration of the nursing program. Improving student assessment practices and ultimately student learning outcomes has been a goal and part of the long term strategic plan at the nursing program.

During the focus groups, the faculty expressed a concern that the current practices at the nursing program are not working for all students and discussed their ongoing frustration with the status quo. The nursing program is in the process of a curriculum change; therefore, this FPD program will assist the faculty when designing new student learning outcomes and student assessment methods.

Duration. Desimone (2002) recommended that the content be revisited over the course of a semester and should be at least 20 hours in length. This workshop will be over the course of three days (21 hours) and faculty will be asked to view two videos as well (1 hour). Also, there will be additional time allotted by having four monthly one hour lunch sessions which will total an additional four (4) hours in length. Therefore the duration of this entire program is 25 hours of face-to-face instruction. Additionally, faculty will be required to read the needed peer review articles prior to each luncheon session but this time is not included in the program duration.

Collaborative participation. Desimone (2009) stressed that learning is best when groups of teachers with similar backgrounds or needs work collaboratively. This idea will be an important aspect of the planned workshop. Faculty will be divided into three small groups of five participants based upon course teaching responsibilities and areas of expertise to allow them to collaboratively re-work many of the MCQ examinations and begin to design rubrics and other strategies to assess written assignments or observed skills.

In numerous articles, Mertler (2003, 2005, 2009) reported the impact FPD can have on faculty's understanding of the assessment of student learning outcomes. Using a

parallel mixed method design, Mertler (2009) studied a small group of professional in-service educators who attended a two week workshop about the assessment of student learning outcomes. None of the educators had any formal training about classroom assessment and self-reported that their knowledge of the topic was limited. *The Assessment Literacy Inventory* was administered as a pretest before this intensive workshop. The participants were also asked to keep daily journals about their perceptions of assessment and reflect upon each workshop session. Following the workshop, the *Assessment Literacy Inventory* was repeated and the results were significant. The overall mean improved in all areas of the inventory and each individual participant's score increased by 6 to 11 points. Mertler's research supported the idea that professional development can be an effective method to address faculty learning needs about student assessment.

Authentic Assessment

There is ongoing debate and controversy about the meaning of authentic assessment and the concept is often confused with performance-based assessment or formative assessment (Frey & Smith, 2007). One of the first educators to define authentic assessment is Wiggins (1989) who wrote, "Assessment is authentic when we directly examine student performance on worthy, intellectual tasks" (p. 1). Others have described this concept as an assessment that mirrors or reflects real world applications (Frey & Smith, 2012; Gulikers, Bastianens, & Kischner, 2004; Mueller, 2012; Pilcher, 2009; Stiggins, 2008).

Nursing is a practice discipline; therefore assessments that measure or provide information about how students will perform in an actual clinical setting are needed. One definition of authentic assessment that is applicable to this nursing program is the one devised by Gulikers, Bastiaens, and Kirschner (2004) who designed a five dimensional framework for authentic student assessment. They defined it as "an assessment requiring students to use the same competencies or combinations of knowledge, skills, and attitudes that they need to apply …in professional life" (p. 69). In a survey of 102 university faculty in England, Whitelock and Cross (2011) determined that while 25% of participants had not heard of the term authentic assessment, 80% believed that methods to assess student outcomes should be real world, relevant, and linked to learning outcomes and objectives.

Following each clinical day, students receive feedback about their performance from nursing faculty. Over the course of each clinical rotation students are required to submit clinical assignments with real world applications. These include patient care and health teaching plans, process recordings, and reviews of evidence-based practice. All of these assignments are graded as pass or fail and are not part of the overall theory grade. Students often expend a great deal of time and effort on these assignments and comment on course-end evaluations that these assignments should be part of their theory grade. These assignments reflect what nurses actually do in a clinical setting and are an excellent example of an authentic assessment; however, students who barely met the objectives receive the same passing grade as students whose work is exceptional. Currently there are no rubrics or other methods to evaluate whether students have met the

objectives. Therefore, during the workshop, one session will be devoted to the creation of a rubric for one of the clinical assignments in each course.

Assessment experts maintain that students need to be evaluated based upon a variety of assessments and the authenticity of each assessment will determine the ability of students to use higher order thinking. Using only teacher-developed MCQ examinations hinders faculty's ability to evaluate critical thinking and clinical judgment and limits opportunities for the development of metacognition or reflection on the part of students (Ambrose, Bridges, Di Pietro, Lovett, & Norman, 2010; Anderson & Krathwohl, 2001).

Formative Assessment

Methods to formatively assess students in higher educational settings have been studied extensively (Jenkins, 2010; Knight, 2006; Mayya, 2010; Miller, 2009; Taras, 2008; Williams, 2011). Baroudi's (2007) definition of formative assessment described it as a process that teachers use to provide feedback to students and determine students' ongoing learning needs. Stiggins (2008) has termed this process assessment for learning rather than the traditional approach which is the assessment of learning. Research has confirmed that formative assessment is often not a well-established practice in many nursing education programs (Duers & Brown, 2009; Koh, 2010; Mayya, 2010; Miller, 2012).

The narratives of the faculty validate the need for ongoing formative assessment. During the focus groups the faculty expressed frustration at the amount of content that needed to be delivered and the structure and format of the curriculum which mandates

that new content be covered each week leaving little time or opportunities to use formative assessment methods. They also provided examples of students who are clinically proficient yet were unable to pass MCQ examinations.

Formative assessments assist the faculty to determine if learning is taking place and afford students the opportunity for self-reflection and self-regulated learning (Jenkins, 2010; Nicol, 2009). Most educators surmised (Williams, 2011; Willis, 2007) real formative assessment is an ongoing process rather than simply individual assessments which are labeled as formative. This process provides ongoing feedback to the student and the goal is an improvement in learning outcomes (Williams, 2011). Willis (2007) described formative assessment as a process that changes students from passive recipients of learning to active participants who accept responsibility for their own learning needs.

Currently, due to the nature of the curriculum, the class schedules, and the amount of content to be delivered, faculty have been unable to provide ongoing and effective feedback to students. Most of the faculty interviewed believed that each unit examination was formative in nature and that effective feedback was provided during each test review. Dunn and Mulvernon (2009) stressed that individual graded examinations can be a formative assessment as they can indicate where the student is on the learning continuum. However, they noted if students have already attained a grade and the grade is high stakes in nature, it is by definition summative since as it is "designed to determine a student's academic development after a set unit of material" (p. 3). Mc Manus (2008) discussed the attributes of effective formative assessment and contended that an assessment is

formative in nature if it used during the instructional period to guide student learning and noted that there is "no such thing as a formative test" (p. 3).

Other authors debated the need for ungraded formative assessments and supported using graded classroom or online quizzes as a formative assessment (Bonnel & Boehm, 2011; Dobson, 2013; Kibble, 2007). As previously noted, some of the faculty believed thought that reviewing each unit test addresses these themes adequately since the program has an excellent initial NCLEX-RN pass rate. However, others were concerned about the high attrition rate and the inability of students who excelled clinically to be successful in the theory portion of each course. Since no MCQ examination can assess metacognition or the ability to create new knowledge (Knight, 2006), and these qualities are important to be an effective nurse who is able to demonstrate clinical judgment, it is vital for the faculty to find methods to adequately assess students in these areas. Boud and Falcikov (2005) maintain that effective formative assessments create life-long learners who are able to use reflection to modify and regulate their own learning.

As nursing experts call for a radical transformation in nursing education, innovative methods of formative assessment have been designed and used successfully in many nursing programs (Billings & Halstead, 2009; Bristol & Zerwekh, 2011). As part of this planned program, faculty will have the opportunity to review formative methods that have been supported in the literature including but not limited to online discussion boards, e-portfolios, audience response systems, classroom assessment techniques, and reflective journals (Bradshaw & Lowenstein, 2010).

Test Review Best Practices

The best practices for item writing, test construction, and item analysis were reviewed in depth for Section 1 of this paper. However, based upon the thematic analysis in Section 2 it is clear that the current method for reviewing examinations is not effective. Since the majority of the faculty are committed to the current assessment practice, it is imperative that the faculty conducted test reviews be a method to provide feedback and guide future student learning.

There is limited research into this aspect of assessment (Weimer, 2012) and many of the articles about this practice are anecdotal in nature (Golding, 2010; Ingram & Nelson, 2006; Poorman & Mastorovich, 2008). Rather than simply providing students with the correct answers and rationales during a final exam testing mathematical calculation for word problems, Golding (2010) administered the exam as two steps. During the first step, students took the test in a traditional format and provided answers to questions. In the second step, the students retook the test through the eyes of the teacher as a self-assessment. Students were provided with the key and told to evaluate their answers based upon the key. If the answer was incorrect, the students were able to immediately see alternative solutions and evaluate misconceptions. Even if they answered a problem correctly, they needed to provide rationales for their calculations. The tests were then handed in for grading. Student feedback was overwhelmingly positive. Although the students received a grade on the test, Golding submitted it was formative in nature as they were able to receive immediate feedback and have the opportunity to correct misconceptions and remediate.

Immediate feedback assessment techniques have also been used to provide instantaneous answers to students during MCQ examinations. Numerous studies (Epstein, et al. 2010) support that college students retained material and learning was enhanced using this innovative technology. These scoring sheets are similar to a lottery scratch off sheet and enable students to keep on choosing answers until they select the correct response. Points are awarded based upon the number of attempts. Students leave the test knowing the correct answer and therefore have the opportunity for reflection. This method also eliminates the need for traditional test reviews and enhances test security.

One method that addresses metacognition and problem based learning that has proven successful in nursing and other higher educational settings is the use of collaborative testing. There has been considerable research into the use of collaborative testing as a method to enhance teamwork and critical thinking, two needed attributes in nurses (Wiggs, 2011). In some studies collaborative testing was used for grading purposes (Bloom, 2009; Gallagher, 2009; Giuliodoni, Lujan & Di Carlo, 2007) while in others it was a method for a posttest review to enhance metacognition as students needed to debate and validate their answers with their peers (Centrella-Nigro, 2011; Poorman & Mastorovich, 2008; Stark, 2006). The evaluation of both methods has been mixed. While research has indicated that either method was a positive experience for students as it enhanced learning and remediation (Gallagher, 2009; Sandal, 2009); Molsbee , 2013) found that if the collaborative test was graded it tended to help lower performing students

attain passing test grades and led to higher attrition of these students in courses that did not use collaborative testing.

Both faculty and students of the nursing program have expressed frustration and believed the entire test review process needs to be changed. Weimer (2012) asked an important question about test preparation and reviews. In a webinar titled "*How can I make my Exams more about Learning and Less about Grades?* (2012), she offered numerous suggestions for conducting what she termed test debriefing. This 20 minute webinar and an introduction to the concept of collaborative testing will be part of the workshop in a session about test reviews. Faculty have been resistant to trialing this method in the past due to concerns about test integrity and cheating. Much of the literature alluded to this aspect (Sandal, 2009) and suggestions to ensure test security will be included during the workshop session about collaborative testing.

Summary of Literature Review

Using workshops and educational sessions to promote faculty development is supported by educational research. For years, these programs focused on faculty satisfaction; there is now a movement by the professional development community to tie effective faculty development to enhanced student learning outcomes (Desimone, 2009). The faculty at this nursing program have expressed a need for professional development in the area of evidence-based assessment of student learning strategies. To provide effective assessment of student learning, assessment experts noted that authentic and formative methods need to be incorporated into curricula to ensure that all types of learners have the opportunity to demonstrate the attainment of learning outcomes.

Implementation

The faculty development workshop has the support of both the administration of the nursing program and the faculty. The schedule at this program is unique; faculty return to school a full two weeks before students which will allow them to actively participate in this program without the added constraints of classroom, laboratory, or clinical teaching responsibilities. Unlike other colleges, faculty are employees of the affiliating hospital, therefore they work a traditional 40 hour work week. The workshop will be scheduled on Wednesday, Thursday, and Friday during the first week of faculty obligation to give the faculty office time to prepare for the upcoming semester during the second week before students arrive on campus. Following this initial workshop, the faculty will participate in four monthly lunch and learn sessions. This workshop will replace the traditional program offered each semester by the Faculty Professional Development Committee. The Director of Nursing Education has budgeted $500.00 for food, resources, and incidentals for this workshop (Director of Nursing Education personal communication, May 2013).

As an incentive for participation in this workshop, I will also submit the required paper work and needed documentation for contact hours through the state nurses association. To maintain licensure and nursing specialty certifications, nurses need varying amounts of demonstrated professional development sessions with contact hours. The hospital which employs the faculty are contact providers therefore this will be at no addition cost to the faculty or the nursing program. It will make this process longer but I

believe that the faculty will appreciate the addition of 12 to 15 contact units to their licensure and certification needs.

Workshops

Workshops and problem based learning are designed to emphasize participant discussion and are used to facilitate the development of skills and competencies (Laureate Education, 2010). Over the course of two days, the faculty will participate in the following workshops and subsequent lunch and learns sessions which are fully explained in Appendix A:

1. Session 1:The How and Why of Test Blueprints (4 hours)
2. Session 2: Good, Better, Best: Evaluating our own Multiple-Choice Questions (3 hours)
3. Session 3: Reflections (1 hour)
4. Session 4:Abracadabra: A Simple Way to Change Multiple-Choice Questions to Assess Higher Order Thinking (4 hours)
5. Session 5: Test Reviews: Is There a Better Way? (3.5 hours)
6. Session 6: Reflections (1/2 hour)
7. Session 7:Assessment Review: Authentic and Formative Methods (2 hours)
8. Session 8: Rubrics: Building Blocks for Success (1.5 hours)
9. Reflections (1/2 hour)

Following the workshop, faculty will have the opportunity to participate in four monthly lunch and learn sessions:

1. Collaborative Testing, (1 hour)

2. Let's bring Our Program into the 21st Century: Online Discussions and other E-Learning Opportunities, (1 hour)
3. Using Technology to Assist with the Assessment of Student Learning, (1 hour)
4. Summative Evaluation of Faculty Professional Development Program: Reflection upon the questions: Where are we now and where are we headed? (1 hour)

Needed Resources

Environmental. To implement this workshop, a smart classroom that has a projector and web access will be used to allow faculty to interact with online resources including Webinars, YouTube® videos, and examples of online discussions, blogs, and other E-Learning strategies. The proposed room is the main skills laboratory in the nursing program. It is equipped with 24 laptops and network capability. The positive aspects of using this area are cost and ease of scheduling. The room is already assigned to the nursing program. The negative aspect is the configuration; it contains 30 desks which will need to be moved to the back of the room and replaced with four small tables to allow participants to work collaboratively. These will be requested from the facilities department for no charge.

Learning Aides/Instructional methods

All participants will be provided with a laptop and USB flash drive device. The laptops will allow participants to have access to the internet and word processing. The flash drive will enable participants to save their work and will also include the session handouts, worksheets, and copies of PowerPoint slides. Reference materials including books, peer reviewed articles, and notebooks will be available on all tables and a full

listing of instructional aides is found in Appendix A. Faulty will also view PowerPoint presentations, webinars, videos, and observe asynchronous threaded discussions. I will be the primary instructor and seasoned faculty with a demonstrated proficiency in the assessment of student learning will be the discussion leaders and facilitators at each table. The lunch and learn sessions will take place in the faculty lounge and are designed as journal articles reviews and group discussion. I have selected two peer reviewed nursing education articles (see Appendix A) for the first two months, an one hour online webinar for the third month, and an evaluative, reflective session is planned for the final session.

Existing Supports

The Director of Nursing Education and the Faculty Professional Development Committee fully support this workshop (Curriculum Committee Minutes, March 2013). The Director has assured me that the faculty schedule will not conflict to enable all faculty to attend. She has budgeted $500.00 for a continental breakfast and lunch and/or on both days and to cover the costs of the flash drives and instructional aides. Refreshments and food will also be provided during the luncheon sessions.

Nursing administration encourages a nurturing environment; therefore, I believe that the faculty will be supportive and receptive to learning in a workshop setting and to participate in the subsequent lunch and learn activities. It is our practice to eat together in the lounge each day and to share family and personal anecdotes. We consider ourselves friends and are cheerleaders for our colleagues, open to learning from each other, and we value the insights and ideas of our more seasoned and experienced members.

Potential Barriers

As noted in the literature (Desimone, 2009; 2011) a major obstacle to effective faculty professional development is time. The planned time frame is during the initial week of faculty fall obligation which means that faculty will be concerned about preparing for the upcoming semester, making clinical schedules and arrangements, ensuring the availability of supplies, answering emails, attending course and curriculum meetings, and meeting with new students or advisees. Planning times are traditionally very hectic; however, these sessions are also the time when the Faculty Professional Development Committee has always scheduled a one-day session each semester and the college's Center for Enrichment of Learning and Teaching (CELT) conducts educational opportunities; therefore, the faculty is accustomed to participating in FPD during this timeframe.

The second potential barrier is the inherent design of the current curriculum at the nursing program which may impact opportunities for the eventual implementation of the assessment strategies discussed at the workshop. Due to class size and environmental constraints, engaging students and using active learning strategies has been difficult. Many of the authentic and formative assessments that will be considered are more effective with small class sizes which allow students and faculty to actively discuss course content (Nilson, 2010; Suskie, 2009). For example, using Socratic questioning, portfolios, or asynchronous online discussion for formative assessment requires a change in teaching methods and additional time commitments from both faculty and students.

Roles and Responsibilities

The workshop was designed using Vella's Seven Design Steps (2008) which has delineated clearly defined roles and responsibilities for the instructor and the students. The Design Steps outlined 12 principles for effective adult learning and stressed that learners will view the workshop as successful when the content is relevant and immediately useful to them. She noted this will only occur in a respectful environment that allows all voices to be heard. This workshop has been designed to facilitate discussion and give the faculty information that will be pertinent, useful, and applicable.

I will be the primary instructor/facilitator for this workshop. According to Vella (1994) my role and responsibility will be to maintain a safe, respectful environment, promote open dialogue, and engage the participants in doing "significant work" (p. 3). All of the participants fit the description of a "great learner" as defined by Luke (2011) and their responsibilities will be to bring an open mind and eagerness to learn to this workshop. Luke described a great learner as "one who excels in their chosen field through commitment to excellence in learning and performance. Such a learner seeks to improve performance or gain realizations consistently through deliberate practice" (para. 3). The participants are all expert nurses and are committed to improving the teaching/learning process at the nursing program.

Program Evaluation

Program Outcomes

Following participation in this three day workshop and four monthly lunch and learn sessions about best practices for the assessment of student learning outcomes, nursing faculty will have:

1. Designed course specific test blueprints which delineate Bloom's revised taxonomy, course objectives, and the NCLEX-RN test plan.
2. Used a rubric to evaluate faculty-developed MCQ examinations.
3. Changed five to ten MCQs written at the remembering and understanding levels of Bloom's revised taxonomy to either the applying or analyzing levels.
4. Examined evidence-based methods to conduct pre and post examination reviews.
5. Debated the best way to integrate authentic and formative assessment methods.
6. Created one holistic or analytical rubric to grade a clinical assignment.
7. Discussed the merits of collaborative testing.
8. Defined E-Learning methods for formative and authentic assessment.
9. Analyzed best practices for the assessment of student learning.

Evaluation Method and Justification

Using an outcomes based evaluation plan fits the conceptual framework (Desimone, 2009) and the program planning model (Vella, 2008) which guided the creation of this capstone project. Both authors agree the best way to evaluate the effect of a program is by determining its impact on the individual participant, the educational community, and ultimately the student learning outcomes. By definition, outcomes based

evaluation strategies measure a change in knowledge, attitudes, skills, and behaviors (Brooks, Marsh, Wilcox, & Cohen, 2011) of the participants and the resultant change in practice at the educational setting (Lightner & Benander, 2010).

Vella (2002) explained that to determine if learning has occurred, one needs to look at the indicators of learning and the assigned learning tasks. A learning task is defined as an "open question put to a small group with all of the resources they need to respond" (p. 46). This process will be evaluated immediately following each session by asking reflective questions and gauging if the learning tasks have been completed. For example following Session 1: *The How and Why of Test Blueprints*, each small group will start the process of designing course specific test blueprints and peer review the blueprints of other faculty. The next step in the evaluation process will be to determine if this learning was transferred. This will occur during the final lunch and learn. Using reflective questions (see Appendix A) the faculty will be asked if the test blueprint they created was effective and guided the construction of the faculty made MCQ tests and if this blueprint reflected course content and objectives.

The final step of this outcomes evaluation plan is the most difficult. According to Desimone's Model (2009) simply administering an end of program satisfaction survey, while important, will not be enough to determine the impact or effect of FPD. To determine the long term effectiveness of this program, I have planned a final lunch and learn which will be an opportunity for the faculty to honestly reflect upon the semester and determine if any of the new assessment strategies were incorporated into the curriculum. To establish if student learning outcomes have changed in relation to this

program, student course-end evaluations will be assessed, analyzed, and trended each semester and linked to ongoing curriculum evaluation at the nursing program. Desimone (2009) noted that this is a difficult process but one that should be attempted to critically evaluate the usefulness of the program. As mentioned in Section One of this paper, students at the nursing program typically write lengthy comments about the content and quality of the MCQ examinations. Although anecdotal evidence, a decrease in the amount of comments and a decrease in attrition at the nursing program will be indicative that this FPD program met its outcome. Also, specific questions addressing assessment methods will be added to the course-end evaluations.

By using an outcomes based evaluation approach, I can also assess the negative, positive, intended, and unintended consequences of the program (Sonpal-Valias, 2009). This is an important aspect. While student course end surveys may be positive, the faculty may perceive that the addition of formative or authentic assessments to each course added to an already overextended workload.

Following this workshop, faculty will be asked to complete a Faculty Development Satisfaction Survey (see Appendix A). This survey is already being used at the nursing program and is required by the hospital's Professional Development Department and needed for state nursing association contact hours. This survey was modified to give me immediate feedback about the design and implementation of the program and it will be an opportunity for feedback before the first scheduled lunch session.

Goals of Evaluation

Currently at the nursing program, the assessment and evaluation of student learning outcomes is a priority. The program will be reviewed in 2014 for reaccreditation by the Accreditation Commission for Education in Nursing Incorporated (ACEN, formerly NLNAC) and the college has been actively preparing for a return visit from the Middle States Association of Colleges and Schools. During the initial site visit from Middle States, the reviewers noted that improvement was needed in the college wide assessment of student learning outcomes and overall program evaluation.

Relevant Stakeholders

The nursing program's stakeholders include the students, the college community, area health care agencies, and the residents of the county that support and fund this program. These individuals have an interest in the success of the nursing program and actively support its existence. Stakeholders are represented on the *Nursing Program Advisory Committee*. Membership includes the following individuals and representatives:

- President of the College,
- Chief Executive Officer of the employing hospital,
- Chief Nursing Officer of the hospital,
- The Director of Nursing Education.
- The Program Coordinator,
- Alumni of the nursing program employed as nurses in the area,
- Program director of an area Bachelor's in Nursing program,

- Program director of an area vocational school,
- A physician with an interest in nursing education, and
- Three nursing representatives from affiliating agencies and clinical sites.

The committee meets each semester to discuss pertinent changes within healthcare and suggest ways to improve the quality of the program and its graduates. The profession of nursing is constantly evolving and input from practice partners is imperative to maintain a relevant and vibrant curriculum. Information about this workshop and any impact on student learning outcomes will be verbally reported during this meeting as faculty professional development in the area of student assessment may ultimately decrease student attrition and increase initial NCLEX-RN pass rates.

Social Change Implications

Although this capstone research project was a small locally conducted case study, I believe it may have far reaching implications within the nursing education community. When I initially reviewed the literature, there were few articles or published research about the issues surrounding the use of only one method to assess student learning and assign grades in nursing programs. Most of the information I obtained about nursing program assessment practices were from unpublished Master's theses or Doctoral dissertations. My review of the literature found only one national study into the assessment and grading practices in nursing programs (Oermann, Saewert, Charasika, & Yarbrough, 2009); and no research into using only one method to asses learning and assign grades. Despite sociocultural changes within the United States, nursing programs primarily rely upon teacher made MCQ examinations to assess student learning and

assign grades. This practice has been challenged by assessment experts who agree that fair, unbiased, reliable, relevant, real world, and ethical student assessments need to be designed and utilized in nursing education program to promote a diverse nursing workforce (Benner, Stuphen, Leonard, & Day, 2010; Institute of Medicine, 2010; NLN Board of Governors, 2012, 2013).

While there has been a transformation in teaching and learning methods in higher education as faculty have embraced newer educational theories, in some respects nursing curricula have not evolved and many still resemble programs from earlier decades. (Billings & Halstead, 2009). Benner (2010) summed up the current crises in nursing education assessment practices when she wrote:

> How learning is assessed sends a powerful message about what the profession believes to be important…we found too few means of assessing the student knowledge and skill acquisition necessary for practice and we worry that there is too much focus on strategies to answer multiple choice questions. (p. 221)

The introduction of evidence-based assessment methods derived from current research and reflective of educational theories will be used to promote positive social change within this nursing program. The students are culturally, economically, and socially diverse and primarily non-traditional. Research has confirmed that these students have unique learning needs and styles that require innovative assessment methods (Ackerman & Barger, 2010; Bosher & Pharris, 2009; NLN, 2012). There has been a paradigm shift in nursing education from the exclusive use of behavioral methods to incorporating concepts and approaches from constructivism, brain-based learning,

social cognitivism, social learning theory, feminist and narrative pedagogy, ethnography, and many other diverse learning theories (NLN, 2003, 2005). Ensuring that all domains of learning and all types of learners are assessed accurately and fairly is paramount to meet the demand for nurses in today's rapidly changing and technology driven healthcare system.

Incorporating new and innovative formative and authentic assessment methods into the nursing program may help retain at risk students and add to the estimated one million new nurses needed by 2018 (NLN, 2010). Using authentic assessment methods will enhance student's ability to demonstrate critical thinking, clinical reasoning, and priority setting. By not relying solely on high stakes teacher-developed MCQ examinations to assign grades, students will have increased opportunities for reflection and self-regulated learning. Nursing students today need to be well versed in technology, collaboration, communication, leadership and management, and evidence-based practice (QSEN, 2013). These are skills and competencies that cannot be assessed by teacher-developed MCQ examinations alone.

Maintaining a vibrant and state of the art nursing program is a primary objective of all stakeholders. The nursing program's primary partner is a Magnet designated hospital and its mission incorporates patient safety, quality, evidence-based practice, education, discovery, and innovation (ANCC, 2014). Changing student assessment practice may positively impact initial NCLEX pass rates which will ultimately increase the supply of qualified nurses within the local community who can demonstrate clinical judgment and a focus on patient safety.

Conclusion

I designed a three day professional development workshop followed by four lunch and learn sessions about best practices for the assessment of student learning outcomes. Since the inception of the program, faculty have decided that the best way to assess student learning outcomes and assign grades was by the exclusive use of teacher-developed MCQ examinations. Following an intrinsic case study and categorical analysis it was determined that faculty perceived that this practice:

- Limited the opportunity for formative assessment of student learning ,
- Limited the faculty's ability to assess higher order thinking; and,
- Limited the development of metacognition of the part of students.

Based upon this, a FPD program about evidence-based assessment practices in higher education was designed. The conceptual framework and program model that guided the creation of this workshop are Desimone's (2009) model and Vella's (2002) Seven Design Steps. An outcomes based approach will be used to evaluate the success of this program. Understanding the faculty and student issues related to using only teacher-developed MCQ examinations and providing FPD may ultimately promote social change as it will assist the faculty to utilize fair, equitable, and evidence-based student assessment practices for this program. This may decrease student attrition and increase the supply of registered nurses to meet increasing societal demands.

Section 4: Reflections and Conclusions

Introduction

Since it opened in September 2007, this nursing program has been using faculty developed MCQ examinations as the only method to assess student learning and assign grades. Despite an excellent initial NCLEX-RN pass rate for every graduating class, the faculty expressed concerns about a high student attrition rate and inability to retain clinically proficient students due to failing theory grades. I conducted an exploratory case study and the following questions were asked:

What are the issues related to the use of teacher-developed MCQ examinations as the only method of student assessment and evaluation in the theory component of courses at a suburban associate degree nursing program?

Subquestions included:

1. How has using only teacher-developed MCQ examinations impacted the teaching/learning process?
2. How has using only teacher-developed MCQ examinations impacted the teacher/ student relationship?
3. How are students assessed formatively in the theory component of the nursing courses?

Following categorical analysis, three major themes emerged. Using only one method to assess student learning and assign grades has:

- Limited the opportunity for formative assessment of student learning ,
- Limited the faculty's ability to assess higher order thinking; and,

- Limited the development of metacognition of the part of students.

To address these issues a faculty professional development workshop about the assessment of student learning was designed based upon current educational best practices.

Since the completion of this research, many changes have occurred at the program. The original director of nursing who gave permission for this study has retired and the former program coordinator accepted the directorship. Following that appointment, I interviewed for, and ultimately accepted the position of program coordinator. I am now in a position to facilitate the implementation of needed revisions to the assessment of student learning practices at the nursing program.

In this section, I will explain the strengths and limitations of my capstone research and faculty development project and discuss what I have learned about myself over the course of this lengthy educational journey in my professional life as a nurse, an educator, a researcher, and a scholar/practitioner.

Project Strengths

The primary strength of the Faculty Professional Development Workshop (FPD) is its format. I designed it based upon the most recent evidence for FPD best practices (Desimone, 2009), and used a program planning model (Vella, 2008) that required active learning on the part of participants. Desimone's conceptual framework is evident in every aspect of this workshop. It is content focused, coherent, uses active learning strategies, promotes collaborative participation, and is of sufficient duration to ensure transfer of learning.

I also designed the program based upon my own experiences as a learner participating in FPD. Too often, I have attended conferences, seminars, and conventions when the primary teaching strategy was PowerPoint presentations and there was no opportunity for me to actively participate. I often returned home remembering little of what I learned and placed the handouts in a file only to be forgotten and eventually discarded. I knew that the programs that meant the most to me were the ones that required me to listen intently and do a significant amount of work.

Doyle (2011) described what I was trying to do with this workshop best when he stated "It is the one who does the work who does the learning" (Doyle, 2011, p. 7). This is a tenet of constructivism and the workshop as a learning method evolved from that theory (Brandon & Alf, 2010). This workshop builds upon the faculty's previous knowledge and experiences and enables them to construct new ideas upon that foundation. It also allows faculty to acquire new information using the Five E's of Learning (Bybee, 2009). To meet the learning objectives participants are required to engage, explore, explain, elaborate, and evaluate the content.

Project Weaknesses

While I believe strongly that this workshop has been designed based upon the most recent literature and best practices; ultimately, I have no control over the participants. My colleagues may come to this workshop with an open mind and be excited about participating in the learning tasks I have designed; or, they can attend and decide to be passive recipients. I have tried to address this aspect by creating concrete

learning tasks for each session and making the content relevant for their professional and educational needs.

Another potential weakness is that many aspects of current curriculum of the nursing program are teacher-centered, not student-centered which may make implementation of a variety of formative assessment difficult in some situations. Lecture remains the primary method of teaching for some of the faculty as classes usually contain 75-90 students who are housed in a large lecture hall with stationary chairs. This may leave little opportunity to use small group or problem based methods to formatively assess student learning. Nilson (2010) noted that while a lecture is a good method to convey factual knowledge, it is not the best to promote critical thinking and problem solving skills. Integrating active learning into a lecture setting is difficult but not impossible and some faculty may need guidance in that area. The goal of the workshop is to give the faculty tools to use formative and authentic assessment methods; this can be challenging in a content- laden curriculum with large class sizes.

Potential Alternative Projects

Curriculum change is a priority at the nursing program and currently the faculty are conducting a review of the literature and researching successful program models. An entirely new curriculum is being considered and there are plans to obtain additional laboratory and classroom space once a new building on campus is completed. The director of nursing education in conjunction with all faculty and program stakeholders produced a 5 year strategic plan outlining all of the long and short terms goals for the nursing program.

There are also many other ways to address faculty professional development needs (POD, 2013). Peer coaching is an excellent strategy to establish feedback and reciprocity among professionals (Donner & Wheeler, 2009; Waddell & Dunn, 2005). Peer coaching was first described by Joyce and Showers (1982) and involves teachers giving support and assistance to their peers. In their original research, the authors (1982, 1996) found enhanced transfer of learning and long term retention of new learning.

In addition, peer coaching decreased faculty's feelings of isolation, and increased collegiality. It is non-evaluative in nature and mutually beneficial to both the coach and the learner. Curriculum change is a priority at the nursing program and currently the faculty are conducting a review of the literature and researching successful program models. A new format for course sections is being discussed and there are plans to obtain additional laboratory and classroom space once a new building on campus is completed. The director of nursing education in conjunction with all faculty and program stakeholders produced a 5-year strategic plan outlining all of the long and short terms goals for the nursing program.

Joyce and Showers (1982) described three types of peer coaching and named them mirroring, collaboration, and expert. Since the nursing faculty are experienced educators, expert coaching would be the best format. The expert coach has more expertise or familiarity with a subject. Currently at the program, five faculty have been identified by their peers as experts at student assessment. Similar to mentoring, the expert coach encourages reflection, analysis, support, and professional development.

Reciprocity is enhanced by a mutual commitment and agreement that peer coaching is not evaluative and all discussions are confidential (Waddell & Dunn, 2005).

Peer coaching teams must agree to practice the new skill, support one another in designing and implementing new strategies, and collect data about how this change or new skill impacts student learning outcomes. Both the coach and the learner need to make a time commitment to this process (Joyce & Showers, 1982). The amount and design of coaching sessions vary in the literature (Houston & Weaver, 2008; Mc Gatha, 2008). I believe this method would also be successful to address faculty learning needs in the area of assessment of student learning, however, it is often a lengthy process and not all experts are effective coaches. For this, reason I ultimately decided to use the workshop format.

What was Learned about the Process

I vividly remember during my first semester at Walden reading Brookfield's seminal work (1987) about critical thinking. In the introduction he wrote: "thinking critically, reflecting on the assumptions underlying our and other's ideas and actions, and contemplating alternative ways of thinking and living-is one of the important ways in which we become adults" (p. x). If I may be so daring as to paraphrase him, I believe that thinking critically and reflecting on our underlying assumptions and beliefs is the first and the most important step in becoming a scholar. In the next section I will outline what I learned about scholarship, program development, leadership, and change over the course of this capstone project.

Scholarship

The NLN, The American Association of Colleges of Nursing, and the Canadian Association of Schools of Nursing (CASN) have all defined scholarship in nursing education based upon Boyer's (1990) definition. All of these organizations stressed that scholarship in nursing education is evident by a wide range of activities that include "the generation, validation, synthesis, and/or application of knowledge to advance the teaching, research, and practice of nursing" (CASN, 2013, p. 2). In my opinion, nursing education scholars are those individuals who contribute significantly to the body of nursing education knowledge. Scholars are the experts, the researchers, the theorists, the role models, the mentors, and the leaders of my chosen profession.

Research is an integral part of scholarship, and when I started my capstone project I knew that I was in a good position to conduct a rigorous research project. In my graduate program, I was required to design and implement a pilot research study with two classmates. As a nurse practitioner, I was the primary investigator for a clinically focused descriptive pilot study which was well received and published in a peer reviewed journal (Siegel, 2006). I served for 4 years on the Nursing Research Committee at my employing hospital and taught basic research concepts to nursing students at the diploma and associate degree levels. I am also an article reviewer for two leading refereed nursing journals.

However, writing this capstone project has convinced me that scholarship is about so much more than research alone. It is about critically reading the literature and constantly seeking and evaluating the best evidence. To be a scholar requires one to be

open and value all opinions and world views. A scholar reflects daily upon his or her professional responsibilities and asks how improvement is achieved.

When my proposal was accepted, I was asked by the nursing representative on the hospital's IRB to present my proposal to the annual nursing research day. I hesitated and told her that staff nurses would not be interested in this topic and I did not want to bore them with nursing education problems or research. She would not back down and I relented and did present my topic. After my presentation, I was approached by two faculty members from other nursing programs that were in attendance. They commented positively on my research questions and said they looked forward to hearing me discuss my research findings at the next annual research day.

I was about to leave when five nursing students from another program approached me and shared with me that they felt exactly as my students did when they read or processed teacher-made MCQ examinations. The students said that they were going to share my handout about my proposal with their faculty. It was in that moment that I realized that I was living the scholarship of teaching and learning. I learned that even simple, locally conducted research studies contribute to the overall body of nursing and higher education knowledge and may impact others.

As I was reviewing the literature for this section, I came across multiple definitions of scholars and scholarship. Following an exhaustive review of the literature, Martin, Benjamin, Prosser, and Trigwell (as cited in Mc Kinney, 2004) determined that there were three dimensions of scholarship. They wrote, "The scholarship of teaching is three related activities; engagement with the existing knowledge…self-reflection…on

one's discipline, and public sharing of ideas about teaching" (p.1). As I thought about this definition, I concluded that when I presented my proposal I was fully engaged in the scholarship of teaching and learning. I was presenting the best evidence, reflecting upon it, promoting reflection on the part of other educators and students, and I was sharing my ideas in a public forum.

Project Development and Evaluation

As a nurse and as an educator, I have been involved with program development for many years. I have served on convention planning committees, designed one day educational programs about various clinical topics, and spoken widely at local nursing conferences. Until I began Walden, I had no idea that there were program planning models and conceptual frameworks to guide the design, implementation, and evaluation of educational offerings. I had always relied upon other members of the committee with extensive experience and never before questioned the planning processes or researched other options. When I designed a class for professional development I essentially decided upon the content, wrote objectives, created the presentation, and then handed out evaluation surveys. I had considered program planning very different from curriculum development and now I appreciate the similarities between both processes.

Using research and best practices to design an educational program allowed me to see so many areas that needed to be addressed that I had never before considered. Taking the results of my research and deciding the best way to address my peers' learning needs required extensive reflection and research. In the past, I would have done exactly what I said in the first paragraph, nothing more and I would have considered that sufficient.

Now I realize that program planning and evaluation is a separate and distinct discipline with an evolving body of research. Finding Desimone's (2009) conceptual model was eye opening for me and using something so simple and yet so profound as Vella's Seven Design Steps (2008) to create each learning task made the evaluation process seamless.

Leadership and Change

As I previously mentioned, there has been a change in the leadership at the nursing program since I completed the research portion of this project. I am now an administrator and my primary responsibilities are curriculum development and evaluation. The faculty have supported my promotion and have expressed to me that they consider me an "expert" in these areas based upon my actions since starting my doctoral studies.

I have been fortunate in my professional life to have always been considered a leader and change agent by my peers. As a staff nurse, I was always the one who approached administration with questions or concerns and was usually the first one to volunteer to help implement a new strategy or treatment on the nursing unit. The majority of my nursing professional life has been spent in management positions.

Quite honestly, I am not sure if I learned anything new about leadership or change from attending Walden or writing this project. I do believe that this entire process has reinforced what I already knew and caused me to reflect upon ways that I can become a more effective leader and change agent. In my new position, I often attend college-wide meetings where I am the face of the nursing program. My opinions, concerns, and

comments are no longer just my own but reflect the entire nursing program and the nursing profession.

Nursing education methods and role of nursing faculty has been debated extensively over the past decade (Benner, Sutphen, Leonard & Day, 2010; IOM, 2011). Hegarty, Condon, Walsh and Sweeny (2009) discussed the future of nursing education and noted numerous challenges facing academic undergraduate faculty. These included the globalization of healthcare, advancements in technology, the changing patient demographic, and the increased complexities of clinical nursing care. This rapid change means that nurse educators "must be flexible, innovative and willing to re-think traditional methods of nursing education" (Giddens, 2008, p. 8). In other words, all nurse educators must be leaders and change agents.

I am not sure if I would approach this project differently. I had originally considered individual faculty mentoring and coaching to address faculty learning needs and I have outlined the pros and cons of that approach. Ultimately, my research confirms that the program may need to undergo curriculum revision which is a lengthy and arduous process. This process has started and the goal is to complete the new curriculum by the fall 2016 semester.

What was Learned about Self

When I started reflecting upon what I learned about myself over the course of my doctoral studies, I kept thinking about an article I read many years ago when I started a bachelors of nursing program. It was 1980 and my profession was in the initial stages of developing theories, conceptual frameworks, and defining nursing as a science. I was

required to read what has now become a classic work, Carper's (1978) *Fundamental Patterns of Knowing in Nursing.* In the next section I will discuss what I have learned about myself as a scholar, as a project developer, and as a practitioner using Carper's four fundamental patterns of knowing and relating them not only to nursing but to the art and science of teaching. These four patterns of knowing are (a) empirics, (b) esthetics, (c) personal knowledge, and (d) ethics, or the moral and ethical components.

As a Scholar

As a scholar, I use empirics or the science of education every day when I critically read, examine, apply, synthesize, and evaluate the educational literature. Since attending Walden and starting this process, I have learned that I love to do literature reviews! At least once every month I try to review nursing and higher education journals to determine what is new and relevant for my practice. I am comfortable sharing these articles with my peers and colleagues via email links or by posting pertinent articles in the faculty lounge. My friends have laughed and wondered if I ever read *People* magazine.

Since conducting this project, I have also made sure that I always cite any resources I have used when I teach, write reports, or present information for other educators. I know that in the past I did not always do that and now I appreciate how important it is to give credit to sources. Recently, I was asked to serve on a committee to help write a curriculum for a nurse residency program. Before the meeting I emailed links to articles about successful nurse residency programs and wrote a memo to all participants outlining best practices with citations. I believe that this practice has become a part of who I am as an educator.

As a Practitioner

How I practice nursing and teaching, how I develop curriculum and evaluation plans is esthetics or the art of my profession. To me this is the most important part of what I do as an educator. The art of teaching is the way I transfer the research into practice and make an impact on my students, my colleagues, my program, and ultimately my profession as a whole. When I decided to pursue my doctorate, I had many options available to me. The Doctorate in Nursing Practice (DNP) had recently been developed and I could have also attained my Doctorate in Philosophy in Nursing. I contemplated both options but eventually I decided that I *knew* nursing; what I did not fully understand was the art of teaching.

Over the course of my reading for this paper, I found an article about a nurse educator who formed a mentoring relationship with an individual with a Doctorate in Education (Eifler & Veltri, 2010). They wrote, "Nurse educators interested in refining their pedagogy might find it productive to approach education faculty for guidance" (p. 626). As I read this article I saw myself. I had no idea when I started Walden that I would learn from K-12 teachers, or community educators, or professional staff developers employed in businesses; but I did. Writing this project alone made me realize that I missed the daily classroom interactions from individuals who work in all aspects of education. I enjoyed reading discussion boards about technology integration, or writing objectives, or ethical issues in education, and I discovered that nurse educators have so much to learn from our colleagues in other fields.

I have been told by my peers and mentors that I was a natural at teaching. After attending Walden and designing this capstone project I have a greater appreciation for the art of teaching. I always knew that I was effective in my interactions with students and that I loved teaching, but I never fully understood why what I did was successful. I now appreciate the art of teaching and have the tools to refine my practice to meet the ever changing needs of my students and my profession.

As a Project Developer

I believe that this is where what Carper (1978) calls the ethics or the moral component of knowledge comes into play. I would like to discuss my project broadly; to me it is not simply the faculty development workshop I designed in Appendix A. My project is every aspect of what I learned over the past few years. Designing a case study and conducting research at my employing agency required me to always think about my ethical responsibilities to my peers and my program. I was surprised at how it was sometimes difficult to maintain confidentiality about the content of my data analysis. For example, after the focus groups were completed, the entire faculty attended a meeting on campus about assessment practices and it was very challenging not to interject or provide examples of faculty focus group comments at this meeting, as it was very relevant for the discussion.

It was also difficult maintaining confidentially about the project as a whole. After the focus groups, faculty would ask what my research findings were or how I was going to approach this issue at the program. I needed to be vague with my answers and it was

sometimes upsetting that I could not share my findings. Normally, I would have asked my peers to critique my writing but I could not in this situation.

I also learned a great deal about the ethical choices we make each day as teachers. Pope, Green, Johnson and Mitchell (2009) asked 103 educators to describe a situation when they were not sure what was correct or ethical in relation to the assessment of student learning. The authors noted that most moral or ethical dilemmas concerned conflicts between assessment policies and practices and what the teachers' thought would benefit students. This article prompted the realization that the faculty of this nursing program were not alone in their concerns about how their assessment practices impacted student attrition and learning. It also reinforced to me the importance of designing and using assessment methods that are fair, reliable, valid, and based upon best practices.

As a Person

I added this section because it is where I believe I learned the most. Over the course of this project, I had health issues, personal issues, family issues, and employment issues but I never let them get in the way of my goal. This project has taken me longer than I had anticipated and in that respect I am very similar to many of my own students. I also learned that I am more goal oriented and self-directed than I had previously thought.

I discovered that as a teacher I reflect what was said many years ago by the prominent educator, K. Patricia Cross. She wrote, "The task of the excellent teacher is to stimulate apparently ordinary people to unusual effort. The tough problem is not in identifying winners, it is in making winners out of ordinary people" (Cross, 2013, para 3). I see many students who struggle and think that education is beyond their reach. These

individuals inspire me each day as I appreciate the obstacles that so many of them need to overcome to attend college and pursue their dreams. Meeting them and becoming part of their lives has been a gift. I learn from my students every day, and I honestly think that pursuing my own education has given me the opportunity to give see myself in each of my students. I know firsthand what it is like to do homework at midnight or be impacted by an ongoing family situation. I think this entire process has made me stronger and more empathetic.

I need to also express another piece of what I learned about myself as a person. I have always believed in the power of humor and it has been a very healthy coping mechanism for me. I use humor in my classes, in the clinical setting when appropriate, and during this entire process. Comedian Red Skelton once said the following and I believe it also applies to what was reinforced over the course of writing this project.

> I live by this credo: have a little laugh and look around you for happiness instead of sadness. Laughter has always brought me out of unhappy situations. Even in your darkest moments, you can usually find something to laugh about if you try hard enough. (Skelton, 2015, para. 2)

Laughter has helped me overcome rewrites, editing, committee comments, lost references, and so many other unanticipated problems. I have tried to remain grounded and focused on the end point. It was not an easy process and there were days that it was very hard to laugh and on those days I reached out and found the laughter and silliness that I needed in my family, my colleagues, and my students.

Importance of this Study

Implications and Applications

Leading educational theorists support the idea that adults learn in multiple ways (Merriam, Caffarella & Baumgartner, 2007); therefore, methods to assess student learning need to be diverse as well (Chappius, Stiggins, Chappius & Arter, 2012). As this case study has shown, using only one method to assess student learning impacts the ability of students to effectively demonstrate understanding of course content. Educational and nursing leaders have noted that today's graduate nurse needs to function in a highly technical and fast paced environment. Nurses are now major determinants of best practices and are required to use the most recent evidence in every patient care interaction (Candela & Bowles, 2008). A multiple-choice question examination can only tell educators so much about how graduates will react in clinical situations.

The results of my case study support the need for nursing educators to be sensitive to the diverse learning needs of all students. Gone are the days when all nursing students were 18 years old, middle class, and Caucasian. Today's students encompass every ethnic, cultural, age, gender, and socioeconomic background and come to undergraduate nursing programs with varied life experiences (Ackerman-Barger, 2010). Therefore, methods to assess if student outcomes have been met need to be diverse and afford all students an equal opportunity to demonstrate learning.

Implications for Future Research

Further research into student assessment is needed in all areas of higher education. The evidence is clear that faculty often are not assessment experts (Frey,

Schmitt & Allen, 2012; Popham, 2009). This study looked at the practice of using teacher-developed MCQ examinations as the only method to assess student learning and assign grades. It did not attempt to correlate this practice to any other areas of the curriculum. Assessment practices are difficult to isolate from other practices at a college and rigorous studies that attempt to quantify and/or correlate assessment practices to student learning styles, attrition, and other issues are still needed.

There is also a need for ongoing research into how students perceive faculty assessment practices. Do we as faculty help or hinder students from meeting educational outcomes by how we ultimately assess and evaluate their learning? If I decide to continue research into assessment practices I would focus my next study on the student issues related to faculty assessment practices.

This was also a small study conducted at one nursing program. Research is needed into assessment practices in nursing programs nationally and globally. Much of the information about assessment practices in nursing programs is anecdotal in nature and there have been few attempts to correlate assessment practices to success on the licensing examination.

Conclusion

For the past six years, my personal and professional life has focused on my doctoral journey at Walden University. I learned a great deal about myself as a scholar, educational practitioner, change agent, researcher, and project designer. Conducting "backyard research" (Glesne, 2011, p. 41) was difficult but valuable. It opened my eyes to problems and caused me to reflect upon my own practice. It also gave me a deeper

appreciation for the tireless commitment of the faculty and administration at my nursing program. Understanding the teacher and student issues related to using only teacher-developed MCQ examinations may promote social change to help the faculty determine the best teaching and learning and assessment practices for this program.

While this paper discussed one aspect of the nursing program, and made suggestions for ongoing development and improvement, I want readers to know that this program is highly respected with the state. It has consistently maintained an initial NCLEX-RN pass rate that exceeds the national average and following graduation over 60% of its graduates immediately pursue a Bachelors or Master's in nursing. The faculty are involved in all aspects of the curriculum and maintain a vibrant systematic plan of evaluation to ensure that all program outcomes are met.

Nursing education is undergoing a paradigm shift. For years, it was enough for colleges of nursing to produce graduates that had excellent hands-on skills and competencies. Now, the focus is on collaborative practice, technology integration, safety, critical thinking, and clinical judgment. This capstone project highlighted the need for a change in the assessment of student learning practices at one small, suburban associate degree program to ensure that graduates are well prepared to practice in this rapidly evolving healthcare environment.

References

Accrediting Commission for Education in Nursing. (2013). Accreditation Manual. Retrieved from http://acenursing.org/accreditation-manual/

Ackerman-Barger, P. W. (2010). Embracing multiculturalism in nursing education learning environments. *Journal of Nursing Education, 49* (12), 677-682. doi: 10.3928/01484834-20100630-03

Amaro, D. J., Abriam-Yago, K., & Yoder, M. (2006). Perceived barriers for ethnically diverse students in nursing programs. *Journal of Nursing Education, 45*(7), 247-254.

Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., & Norman, M. K. (2010). *How Learning Work: 7 Research Based Principles for Smart Teaching*. San Francisco, CA: Josey Bass.

Anderson, L. W.,& Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

American Nurses Association. (2010). *Nursing Professional Development: Scope and Standards of Practice.* Washington D. C. Author.

American Nurses Credentialing Center. (2014). Magnet recognition program overview. Retrieved from http://www.nursecredentialing.org/Magnet/ProgramOverview

Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers*. San Francisco: Josey Bass.

Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher, 36*(5), 258-267. doi: 10.3102/0013189X07306523

Baroudi, Z. M. (2007). Formative assessment: Definition, elements and role in instructional practices. *Post-Script: Postgraduate Journal of Educational Research, 8* (1), 37-48. Retrieved from http://www.edfac.unimelb.edu.au/research/resources/student_res/postscriptfiles/vol8/Ziad_Baroudi.pdf

Benefiel, D. (2011). The story of nursing licensure. *Nurse Educator, 36*(1), 16-20. doi: 10.1097/NNE.0b013e3182001e82

Benner, P., Sutphen, M., Leonard, V., & Day, L. (2010). *Educating nurses: A call for radical transformation*. San Francisco, CA: Jossey-Bass.

Bible, L., Simkin, M. G., & Kuechler, W. L. (2008). Using multiple-choice tests to evaluate students' understanding of accounting. *Accounting Education: an international journal,17*, S55-S68. doi: 10.1080/0963928080200949.

Billings, D. M., & Halstead, J. A. (2009). *Teaching and learning in nursing: A guide for faculty* (3rd ed.). St. Louis, MO: Saunders Elsevier.

Birman, B. F., Desimone, L., Porter, A. L., & Garet, M. S. (2000). Designing professional development that works. *Educational Leadership, 57*(8)28-33. Retrieved from http://cemse.uchicago.edu/computerscience/OS4CS/landscapestudy/resources/Birman-Desimone-Porter-and-Garet-2000.pdf

Bloom B. S. (1956). *Taxonomy of educational objectives, Handbook I: The cognitive domain.* New York: David McKay Co Inc.

Bonnel, W., & Boehm, H. (2011). Improving feedback to students online: Teaching tips from experienced faculty. *Journal of Continuing Education in Nursing, 42*(11), 503. doi: 10.3928/00220124-20110715-02

Bosher, S. D., & Bowles, M. (2008). The effects of linguistic modifications on ESL students' comprehension of nursing course test items. *Nursing Education Perspectives,29*(3), 165-172. doi: http://dx.doi.org/10.1043/1094-2831(2008)29[165:TEOLMO]2.0.CO;2

Bosher, S. D., & Pharris, M. D. (2009). *Transforming nursing education: The culturally inclusive environment*. New York: Springer Publishing.

Boud, D., & Falchikov, N. (2005). Redesigning assessment for learning beyond higher education. *Research and Development in Higher Education, 28,* 34-41. Retrieved from http://conference.herdsa.org.au/2005/pdf/refereed/paper_398.pdf

Boyer, E. (1990). *Scholarship reconsidered: Priorities for the professoriate*. Princeton, NJ: The Carnegie Foundation for the Advancement of Teaching.

Bradshaw, M., & Lowenstein, A. (Eds.). (2010). *Innovative teaching strategies in nursing and related health professions* (3rd ed.). Boston: Jones & Bartlett Publishers.

Brady, A. (2005). Assessment of learning with multiple-choice questions. Nurse Education in Practice,5(4), 238-242. doi: 10.1016/j.nepr.2004.12.005

Brandon, A. F., & All, A. C. (2010). Constructivism theory analysis and application to curricula. *Nursing Education Perspectives, 31*(2), 89-92. doi: 10.1043/1536-5026-31.2.89

Bristol, T. J., & Zerwekh, J. (Eds.). (2011). Philadelphia: *Essentials of e-learning for nurse educators*. FA Davis Company.

Brookfield, S. D. (1987). *Developing critical thinkers: Challenging adults to explore alternative ways of thinking and acting*. San Francisco: Jossey-Bass.Brooks, D. C., Marsh, L., Wilcox, K., & Cohen, B. (2011). Beyond satisfaction: Toward an outcomes-based, procedural model of faculty development program evaluation. *The Journal of Faculty Development, 25*(3), 5-12. http://www.editlib.org/p/87260.

Brown, J. F. (2008). Developing an English as a second language program for foreign born nursing students at a historical Black university in the United States. *Journal of Transcultural Nursing, 19*, 184-191. doi: 10.1177/1043659607312973.

Butler, A. C., & Roedinger, H. L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice question tests. *Memory and Cognition,36*(13). doi: 10.3758/MC.36.3.604

Bureau of Labor Statistics (2010). Occupational outlook handbook: 2010-2011. Washington, DC: US Department of Labor. Retrieved from http://www.bis.gov/oco/home.htm

Bybee, R. W. (2009). The BSCS 5E instructional model and 21st century skills. National Academies Board on Science Education, Washington, DC. Retrieved from. http://itsisu.concord.org/share/Bybee_21st_Century_Paper.pdf

Caffarella, R. S. (2002). *Planning programs for adult learners.* San Francisco CA: Jossey Bass.

Canadian Association of Schools of Nursing. (2013). Scholarship in nursing. Retrieved from http://casn.ca/wp-content/uploads/2014/10/ScholarshipinNursingNov2013ENFINALmm.pdf

Candela, L., & Bowles, C. (2008). Recent RN graduates perceptions of educational preparation. *Nursing Education Perspectives, 29*(5), 266-271. doi: 10.1043/1536-5026(2008)029[0266:RRGPOE]2.0.CO;2Candela, L., Gutierrez, A., & Keating, S. (2013). A national survey examining the professional work life of today's nursing faculty. *Nurse Education Today, 33*(8):853-9. doi: 10.1016/j.nedt.2012.10.004

Carper, B. A. (1978). Fundamental patterns of knowing in nursing. *Advances in Nursing Science, 1*(1), 13-24. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/110216

Centrella-Nigro, A. M. (2011). Collaborative testing as posttest review. *Nursing Education Perspectives, 33*(5), 340-341.

Chappius, J., Stiggins, R. J., Chappius, S., & Arter, J.A. (2012). *Classroom Assessment for Student Learning: Doing it Right, Using it Well* (2nd ed.). Upper Saddle River, NJ: Pearson.

Clapper, T. (May 5, 2013). Metacognition: Are your learners really thinking about content? [The Evolution Blog]. Retrieved from http://www.evolllution.com/program_planning/metacognition-are-your-learners-

really-thinking-about-the-content/

Clarke, J. & Dede, C. (2010). Assessment, technology and change. *Journal of Research on Technology in Education, 42* (3), 309-328. Retrieved from http://nrs.harvard.edu/urn-3:HUL.InstRepos:4900838

Clifton, S. L., & Schriner, C. L. (2010). Assessing the quality of multiple choice examinations. *Nurse Educator, 35*(1), 12-16.

. doi:1010.1097/NNE.0b013e3181c41fa3

Colorado State University. (2011). Advantages and disadvantages of the survey method. Retrieved from http://writing.colostate.edu/guides/research/survey/com2d1.cfm

Costanzo, R. D. (2009). *The function and use of the textbook in an undergraduate nursing program*. (Doctoral Dissertation). Retrieved from Proquest

Cox, T. & Mayorga, M. G. (2010). Perceptions of differences in components of faculty professional development. *Academic Leadership: The Online Journal, 8*(4), Retrieved from http://contentcat.fhsu.edu/cdm/ref/collection/p15732coll4/id/537

Creswell, J. W. (2008). *Educational research: planning, conducting, and evaluating quantitative and qualitative research* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Cross, K. J. (2000). *Cognitive levels of multiple-choice items on teacher-made tests in nursing education*. (Doctoral Dissertation). Retrieved from Proquest (AA19992186).

Cross, K. P. (2013). Quotations. Retrieved on March 8, 2015 from http://www.quotemountain.com/famous_quote_author/k__patricia_cross_famous_quotations/

Cross, K. P. (1981). *Adult as learners: Increasing participation and facilitating learning*. San Francisco: Jossey-Bass.

Del Bueno, D. (2005). A crisis in critical thinking. *Nursing Education Perspectives, 26*(5), 278-282. doi:10.1043/1536-5026(2005)026[0278:ACICT]2.0.CO;2De Palma, J. A. (2008, April). *Evidence-based teaching for clinical and academic educators*. Workshop conducted at the meeting of the New Jersey League for Nursing, Atlantic City, NJ.

De Pew, D.D. (2001). *Validity and reliability in nursing multiple-choice testing and the relationship to NCLEX success: An internet survey*. (Doctoral dissertation). Retrieved from Proquest (3040905).

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis, 24*(2), 81-112, Retrieved from http://www.jstor.org/stable/3594138

Desimone, L. M. (2009). Improving impact studies of teacher professional development: Towards better conceptualization and measures. *Educational Researcher, 38*(3), 181-199. doi: 10.3102/0013189X08331140

Desimone, L. M. (2010-2011). A primer of effective professional development. *Kappan Magazine*, 28-31. Retrieved from http://www.gcisd-

k12.org/cms/lib/TX01000829/Centricity/Domain/78/A_Primer_on_Effective_Professional_Development.pdf

DiBattista, D., & Kurzawa L. (2011). Examination of the quality of multiple-choice items on classroom tests. *The Canadian Journal for the Scholarship of Teaching and Learning, 2*(2), 1-23. Retrieved from http://dx.doi.org/10.5206/cjsotl-rcacea.2011.2.4

Dobson, J. L. (2013). Retrieval practice is an efficient method of enhancing the retention of anatomy and physiology information. *Advances in physiology education, 37*(2), 184-191.doi: 10.1152/advan.00174.2012

Donnelly, T. T., Mc Kiel, E., & Hwang, J. (2009). Factors influencing the performance of English as an additional language nursing students: Instructors' perspectives. *Nursing Inquiry, 16*(3), 201-211.

Donner, G. J., & Wheeler, M. M. (2009). Coaching in nursing: An introduction. International Council of Nurses. Retrieved from http://www.icn.ch/images/stories/documents/pillars/sew/coaching/Coaching_in_Nursing.pdf

Dowling, S. M. (2005). The effects of violating standard item writing principles on tests and students: The consequences of using flawed test items on achievement tests in medical education. *Advances in Health Sciences, 10*, 133-143. doi: 10.1007/s10450-004-4019-5

Doyle, T. (2011). *Learner-centered teaching: Putting the research on learning into practice*. Sterling, VA: Stylus Publications.

Draper, J. & Clark, L. (2007). Impact of continuing professional education on practice: The rhetoric and the reality. *Nurse Education Today, 27*, 515-517. doi: 10:1016/j.net.2007.07.001

Draper, S. W. (2009). Catalytic assessment: Understanding how MCQs and EVS can foster deep learning. *British Journal of Educational Technology, 40*(2), 285-293. doi: 10.1111/j.1467-8535.2008.00920x

Drummond-Young, M., Brown, B., Noesgaard, C., Lunyk-Child, O., Maich, N. M., Mines, C., & Linton, J. (2010). A comprehensive faculty development model for nursing education. *Journal of Professional Nursing, 26*(3), 152-161. doi: http://dx.doi.org/10.1016/j.profnurs.2009.04.004

Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation, 14*(7), 1-11. Retrieved from http://www.pareonline.net/pdf/v14n7.pdf

Duers, L. E., & Brown, N. (2009). An exploration of student nurses' experiences of formative assessment. *Nurse Education Today, 29*(6), 654-659. doi:10.1016/j.nedt.2009.02.007

Ebert-May, D., Derting, T. L., Hodder, J., Momsen, J. L., Long, T. M., & Jardeleza, S. E. (2011). What we say is not what we do: Effective evaluation of faculty professional development programs. *Bioscience, 61*(7). 550-558. doi:10.1525/bio.2011.61.7.9

Eib, B. J., & Miller, P. (2006). Faculty development as community building-An approach

to professional development that supports communities of practice for online teaching. *The International Review of Research in Open and Distance Learning, 7*(2).Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/299/639

Eifler, K. E., & Veltri, L. M. (2010). Drawing from the outside for support in teaching. *The Journal of Nursing Education, 49*(11), 623. doi: 10.3928/01484834-20100630-02

Epstein, M. L., Lazarus, A. D., Calvano, T. B., Matthews, K. A., Hendel, R. A., Epstein, B. B., & Brosvic, G. M. (2010). Immediate feedback assessment technique promotes learning and corrects inaccurate first responses. *The Psychological Record, 52*(2), 5. Retrieved from http://opensiuc.lib.siu.edu/tpr/vol52/iss2/5

Frey, B., & Schmitt, V. (2007). Coming to terms with classroom assessment. *Journal of Advanced Academics, 18*(3). Retrieved from http://tebege.net/ar/Articles/Coming_to_Terms_With_Classroom_Assessment.pdf

Frey, B. B., Schmitt, V. L., & Allen, J. P. (2012). Defining authentic classroom assessment. *Practical Assessment, Research, & Evaluation, 17*(2). Retrieved from http://pareonline.net/pdf/v17n2.pdf

Fulcher, R., & Mullin, C. (2011, March). *A data-driven examination of the impact of associate and bachelor's degree programs and the nation's nursing workforce* (Policy Brief 2011-02PBL). Washington, DC: American Association of Community Colleges.

Gallagher, P. A. (2009). Collaborative essay testing: group work that counts. *International Journal of Nursing Education Scholarship, 6*(1). doi: 10.2202/1548-

923X.1845

Gardner, J. (2005). Barriers influencing the success of racial and ethnic minority students in nursing programs. *Journal of Transcultural Nursing, 16*(2), 155-162. doi: 10.1177/1043659604273546

Giddens, J. (2009). Changing paradigms and challenging assumptions: Redefining quality and NCLEX-RN pass rates. *Journal of Nursing Education, 48*(3), 123-124. doi: 10.3928/01484834-20090301-04

Giddens, J. F. (2008). Achieving diversity in nursing through multicontextual learning environments. *Nursing Outlook, 56*(2), 78-83. doi:10.1016/j.outlook.2007.11.003

Glesne, C. (2011). *Becoming qualitative researchers* (4th ed.). Upper Saddle River, NJ: Pearson.

Golding, T.L., (2010, August). Final fitness and the Louisiana 2 step. *The Teaching Professor*. Retrieved from http://www.magnapubs.com/newsletter/the-teaching-professor/story/5267/

Gulikers, J. T., Bastiaens, T. J., & Kirschner, P. A. (2004). A five dimensional framework for authentic assessment. *Educational Technology Research and Development, 52*(3), Retrieved from http://www.jstor.org/stable/30220391

Haladyna, T. M., Dowling, S. M., & Rodriquez, M. C. (2002). A review of multiple-choice item writing rules for classroom assessment. *Applied Measurement in Education, 15*(3), 309-334. Retrieved from http://www.comp.dit.ie/dgordon/Courses/ILT/ILT0005/AReviewofMultipleChoiceItemWriting.pdf.

Hancock, D.R., & Algozzine, B. (2006). *Doing case study research*. New York: Teachers College Press

Health Resources and Services Administration. (September 2010). *The Registered Nurse Population: Findings From the 2008 National Sample Survey of Registered Nurses. Washington*, DC: U.S. Department of Health and Human Services

Hegarty, J., Walsh, E., Condon, C., & Sweeney, J. (2009). The undergraduate education of nurses: looking to the future. *International Journal of Nursing Education Scholarship, 6*(1). doi: 10.2202/1548-923X.1684,

Herman, J. W. (2010, March). *Be an innovative nurse educator: Strategies that promote classroom and clinical learning*. Workshop conducted at the meeting of the New Jersey League for Nursing, Atlantic City, NJ.

Hochberg, E. D., & Desimone, L. M. (2010). Professional development in the accountability context: Building capacity to achieve standards. *Educational Psychologist, 45*(2), 89-106. doi: 10.1080/00461521003703052

Huston, T., & Weaver, C. L. (2008). Peer coaching: Professional development for experienced faculty. *Innovative Higher Education, 33*(1), 5-20.doi: 10.1007/s10755-007-9061-9

Institute of Medicine. (2011). *The future of nursing: Leading change, advancing health*. Washington, DC: The National Academies Press.

Jeffreys, M. (2012). *Nursing student retention: Understanding the process and making a difference*. New York: Springer

Jenkins, J. O. (2010). A multi-faceted formative assessment approach: better recognizing the learning needs of students. *Assessment & Evaluation in Higher Education, 35*(5), 565-576. doi:10.1080/02602930903243059

Joyce, B., & Showers, B. (1982). The coaching of teaching. Educational Leadership, 53(6), 4-10. Retrieved from http://eric.ed.gov/ERICWebPortal/detail?accno=EJ269889

Junious, D. L., Malecha, A., Tart, K., & Young, A. (2010). Stress and perceived faculty support among foreign-born baccalaureate nursing students. *Journal of Nursing Education, 49*(5), 261-270. doi: 10.3928/01484834-20100217-02

Kalb, K. A. (2008). Core competencies of nurse educators: Inspiring excellence in nurse educator practice. *Nursing Education Perspectives, 29*(4), 217-219. doi: 10.1043/1536-5026(2008)029[0217:CCONEI]2.0.CO;2

Kenney, J. L., Banerjee, P., & Newcombe, E. (2010). Developing and sustaining positive change in faculty technology skills: Lessons learned from an innovative faculty development initiative. *International Journal of Technology in Teaching and Learning, 6*(2), 89-102. Retrieved from http://www.sicet.org/journals/ijttl/issue1002/1_Kenny.pdf

Kibble, J. (2007). Use of unsupervised online quizzes as formative assessment in a medical physiology course: effects of incentives on student participation and performance. *Advances in Physiology Education, 31*(3), 253-260. DOI: 10.1152/advan.00027.2007

Killingsworth, E. E. (2014). Nursing faculty decision making about best practices in test construction, item analysis, and revision (Doctoral dissertation, Retrieved from Proquest, 3577368.)

Kim, S., Spielberg, F., Mauksch, L., Farber, S., Duong, C., Fitch, W., & Green, T. (2009). Comparing narrative and multiple-choice formats in online communication skills assessment. *Medical Education, 43*, 533-541. doi: 10.1111/j.1365-2923.2009.03368x

Knight, P. (2006). The local practices of assessment. *Assessment and Evaluation in Higher Education, 31*(4), 435-452.doi: 10.1080/0202930600679126

Knowles, M. S., Holton, E. F., & Swanson, R. A. (2011). *The adult learner* (7th ed.). New York, NY: Taylor & Francis.

Koh, L. C. (2010). Academic staff perspectives of formative assessment. *Nurse Education in Practice, 10*(4), 205-209. doi:10.1016/j.nepr.2009.08.007

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212-264. Retrieved from http://www.unco.edu/cetl/sir/stating_outcome/documents/Krathwohl.pdf

Lavandera, R., Whalen, D. M., Perkel, L. K., Hackett, V., Molnar, D., Steffey, C., & Harris, J. (2011). Value-added of HESI exam as a predictor of timely first-time RN licensure. *International Journal of Nursing Education Scholarship, 8*(1). doi: 10.2202/1548-923X.2152,

Laureate Education, Inc. (Producer). (2009). Dialogue Education. [DVD]. Baltimore,

MD: Author

Leung, S. F., Mok, E., & Wong, D. (2008). The impact of assessment methods on the learning of nursing students. *Nurse Education Today, 28*, 711-719. doi:10.1016/j.nedt.2007.11.004

Light, G., Calkins, S., Luna, M., & Drane, D. (2009). Assessing the impact of a year-long faculty development program on faculty approaches to teaching. *International Journal of Teaching and Learning in Higher Education, 20*(2), 168-181. Retrieved from http://library.oum.edu.my/oumlib/sites/default/files/file_attachments/odl-resources/4583/ijtlhe202.pdf#page=94 Lightner, R & Benander, R. (2010). Student learning outcomes: Barriers and solutions for faculty development. *Journal of Faculty Development, 24* (2), 34-39. Retrieved from http://www.ucblueash.edu/ltc/JFD242Lightner1.pdf

Lujan, J. (2007). Linguistic and cultural adaptations needs of Mexican-American nursing students related to multiple-choice testing. *Journal of Nursing Education, 47*(7), 327-330.

Luke, D. (March 3, 2011). What is a great learner? Retrieved from http://thelearningnut.blogspot.com/2011/02/what-is-great-learner.html

M. County College. (2013). Student demographics. Retrieved from http://[redacted]

Maheshwari, V. K. (2012). Workshop: An instructional method. Retrieved from http://www.vkmaheshwari.com/WP/?p=385

Malau-Aduli, B.S. (2011). Exploring the experiences and coping strategies of international medical students. *BMC Medical Education,11*. Retrieved from http://web.ebscohost.com.ezp.waldenulibrary.org/ehost/pdfviewer/pdfviewer?vid=3&hid=7&sid=8628ccc9-3d70-4cad-80a8-4949784085c9%40sessionmgr4.Marsh, E. J., Roedinger, H. L., Bjork, R. A., & Bjork, E. L. (2007). The memorial consequences of multiple-choice testing. *Psychonomic Bulletin and Review*, *14*(2), 194-199. Retrieved from . http://psych.wustl.edu/memory/Roddy%20article%20PDF's/Marsh%20et%20al%20(2007)_PBR.pdf

Martinez, M. E. (2006). What is metacognition? *Phi Delta Kappan*, *87*(9), 696-699. Retrieved from http://www.gse.uci.edu/person/martinez_m/docs/mmartinez_metacognition.pdf

Maskey, C. (2009). Cognitive coaching has an exciting place in nursing education. *Teaching and Learning in Nursing 4* (2) 63–65. doi:10.1016/j.teln.2008.05.003.

Masters, J. C., Hulsmeyer, B. S., Pike, M.E., Leichty, K., Miller, M. T., & Verst, A. L. (2001). Assessment of multiple choice questions in selected test banks accompanying textbooks used in nursing education. *Journal of Nursing Education, 40* (1), 25-31

Mayya, S. S. (2010). Formative assessment practice in health science education. The *Internet Journal of Allied Health Sciences and Practice, 8*(4). Retrieved from http://ijahsp.nova.edu/articles/Vol8Num4/Mayya.htm

McDonald, M. E. (2013). *The nurse educator's guide to assessing learning outcomes* (3nd ed.). Boston, MA: Jones and Bartlett Publishers.

McGatha, M. (2008). Levels of engagement in establishing coaching relationships. *Teacher Development, 12* (2), 139-150. doi: 10.1080/13664530802038147

McKinney, K. (2004). What is the scholarship of teaching and learning (SoTL) in higher education? Retrieved from http://sotl.illinoisstate.edu/downloads/pdf/definesotl.pdf

McManus, S. M. (2008). A study of formative assessment and high stakes testing: Issues of student efficacy and teacher views in the mathematics classroom. (Doctoral dissertation) Retrieved from ProQuest (3306610).

McNamara, A., Roat, C., & Kemper, M. (2012). Preparing nurses for the new world order: A faculty development focus. Nursing administration quarterly, 36(3), 253-259. doi: 10.1097/NAQ.0b013e31825889ee

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey Bass.

Merriam, S. B., Caffarella, R. S., & Baumgartner, L. M. (2007). *Learning in adulthood: A* comprehensive guide. John Wiley & Sons.

Mertler, C. A. (2003, October). Preservice versus inservice teachers' assessment literacy: Does classroom experience make a difference. In Annual Meeting of the Mid-Western Educational Research Association, Columbus, OH. Retrieved from http://wps.ablongman.com/wps/media/objects/1530/1567095/tchassesslit.pdf

Mertler, C. A., & Campbell, C. (2005, April). Measuring teachers' knowledge and application of classroom assessment concepts: Development of the Assessment Literacy Inventory. In annual meeting of the American Educational Research Association, Montreal, Quebec, Canada. Retrieved from http://www2.tec.ilstu.edu/students/tec_303/Past_Material/Student_Portfolios/2006_Portfolios/Wixted/ASEE%20Lit%20Review/Assessment%20Literacy%20Inventory.pdf

Mertler, C. A. (2009). Teacher's assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools, 12*(2), 101-113. doi: 10.1177/1365480209105575

Miller, B. (2012). Student learning: The role of self-assessment" (2012). (Master Thesis). Retrieved from Digital Commons.

Molsbee, C. P. (2013). Collaborative testing and mixed results. *Teaching and Learning in Nursing, 8*(1), 22-25.doi:10.1016/j.teln.2012.09.001

Morrison, S., & Walsh Free, K. (2001). Writing multiple-choice test items that promote and measure critical thinking. *Journal of Nursing Education 40*(1), 17-24. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11198905

Mueller, J. (2012). What is authentic assessment? Retrieved from http://jfmueller.faculty.noctrl.edu/toolbox/whatisit.htm

Mundy, M. A., Kupczynski, L., Ellis, J. D., & Salgado, R. L. (2012). Setting the standard for faculty professional development in higher education. *Journal of Academic*

and Business Ethics 5, 1-9. Retrieved from http://www.aabri.com/manuscripts/111041.pdf

NCLEX-RN Test Plan (2013). Retrieved from https://www.ncsbn.org/2013_NCLEX_RN_Test_Plan.pdf

National Council of State Boards of Nursing. (2011). Exam psychometrics. Retrieved from https://www.ncsbn.org/1230.htm

National Council of State Boards of Nursing. (2010, November). *NCSBN Regional NCLEX test writing workshop*. Workshop conducted at the meeting of NJ Associate Degree Faculty, Somerset NJ.

National League for Nursing .(2012). The Fair Test Imperative in Nursing Education. Retrieved from http://www.nln.org/aboutnln/livingdocuments/pdf/nlnvision_4.pdf

National League for Nursing (2010). Global nursing shortage. Retrieved fromhttp://www.nln.org/aboutnln/globaldiversity/index.htm

National League for Nursing. (2009). A commitment to diversity in nursing and nursing education. Retrieved from http://www.nln.org/aboutnln/reflection_dialogue/refl_dial_3.htm

National League for Nursing. (2005). Core competencies of nurse educators with task statements. Retrieved from www.nln.org

National League for Nursing . (May 9, 2005). Transforming Nursing Education. Retrieved from http://www.nln.org/aboutnln/PositionSTatements/transforming052005.pdf

National League for Nursing. (August 23, 2003). Innovation in Nursing Education: A call

for reform. Retrieved from http://www.nln.org/aboutnln/PositionSTatements/innovation082203.pdf

National League for Nursing. (May 18, 2002). Preparation of Nurse Educators . Retrieved from http://www.nln.org/aboutnln/PositionSTatements/preparation051802.pdf

National League for Nursing. (September 19, 2001) Lifelong learning for nurse educators. Retrieved from http://www.nln.org/aboutnln/PositionSTatements/lifelong091901.pdf

National League for Nursing Accrediting Commission. (2008). Standards and criteria. Retrieved from http://www.nlnac.org/manuals/SC2008.htm

National League for Nursing Summit. (September 22, 2011) Special Session

Nedeau-Cayo, R., Laughlin, D., Rus, L., & Hall, J. (2013). Assessment of item writing flaws in multiple choice questions. *Journal for Nurses in Professional Development, 29*(2), 52-57. doi: 10.1097/NND.0b013e318286c2f1

New Jersey League for Nursing. (2005, March). *The ultimate test: developing valid and reliable exams to assess higher-order thinking*. Workshop conducted at the meeting of the New Jersey League for Nursing, Atlantic City, NJ.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practices. *Studies in Higher Education, 31*(2), 199-218. Retrieved from http://tltt.strath.ac.uk/REAP/public/Resources/DN_SHE_Final.pdf

Nilbert, A. T., Young, A., & Britt, R. (2008). The HESI Exit Exam: Progression benchmarks and remediation guide. *CIN: Computers, Informatics, Nursing, 26*(5),

70S-74S. doi: 10.1097/01.NCN.0000336447.32165.70.

Nilson, L. B. (2010). *Teaching at its best: A research-based resource for college instructors* (3rd ed.). San Francisco, CA: Jossey Bass

Notar, C. E., Zuelke, D. C., Wilson, J. D., & Yunker, B. D. (2004). The table of specifications: Insuring accountability in teacher made tests. *Journal of Instructional Psychology, 31*(2), 115-129.

Oermann, M. H., & Gaberson, K. B. (2009). *Evaluation and testing in nursing education* (3rd ed.). New York: Springer Publishing Company.

Oermann, M. H., Saewert, K. J., Charasika, M., & Yarbrough, S. S. (2009). Assessment and grading practices in schools of nursing: National survey findings Part I. *Nursing Education Perspectives*, *30*(5), 274-278.

Ortelli, T. A. (2011). *Test construction for nurse educators*. Workshop conducted at the meeting of the New Jersey League for Nursing, Atlantic City, NJ.

Ortelli, T. A. (2006). Defining the professional responsibilities of academic nurse educators: The results of a national practice analysis. *Nursing Education Perspectives,2*(5), 242-246. Doi:10.1043/1536-5026(2006)27

O'Sullivan, P. S., & Irby, D. M. (2011). Reframing research on faculty development. *Academic Medicine, 86*(4), 421-425. doi:10.1097/ACM.Ob013e31820dc058

Palmer, E. J., & Devitt, P. G. (2007). Assessment of higher order cognitive skills in undergraduate education: Modified essay or multiple-choice questions? *BMC Medical Education, 7*.doi:10.1186/1472-6920-7-49

Patton, M.Q. (2002). *Qualitative Research and Evaluation Methods* (3rd ed.). Thousand

Oaks CA: Sage

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal, 44*(4). 921-958. doi: 10.3102/0002831207308221

Pilcher, J.W. (2009). Using concept maps in a nurse internship program. *Journal for Nurses in Professional Development*, *25*(6), 299-303. doi: 10.1097/NND.0b013e3181c266f3

Poindexter, K., Hagler, D., & Lindell, D. (2015). Designing Authentic Assessment: Strategies for Nurse Educators. *Nurse Educator, 40*(1), 36-40. doi: 10.1097/NNE.0000000000000091

Poorman, S. G., & Mastorovich, M. L. (2008). Using metacognitive strategies to help students learn in pretest and posttest review. *Nurse Educator, 33*(4), 176-180. doi: 10.1097/01.NNE.0000312199.81815.f4

Pope, N., Green, S. K., Johnson, R. L., & Mitchell, M. (2009). Examining teacher ethical dilemmas in classroom assessment. *Teaching and Teacher Education, 25*(5), 778-782.doi:10.1016/j.tate.2008.11.013

Professional and Organizational Development Network in Higher Education. (2013). What is professional development? Retrieved from http://www.podnetwork.org/development.htm

Powers, B. A., & Knapp, T. R. (2010). *Dictionary of nursing theory and research* (4rd ed.). New York: Springer Publishing Company.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice, 48*(4), 4-22. doi: 10.1080/004005840802577536.

Quality and safety education for nurses. (2013). Pre-licensure KSAs. Retrieved from http://www.qsen.org/

Redmond, S. P., Hartigan-Rogers, J. A., & Cobbett, S. (2012). High time for a change: Psychometric analysis of multiple-choice questions in nursing. *International Journal of Nursing Education Scholarship, 9*(1). doi: 10.1515/1548-923x.2487

Rees, B. B. (2006). Can you have both retention and increased pass rates on the NCLEX-RN? *Teaching and Learning in Nursing, 1,* 18-21. doi: 10.1016/j.teln.2006.4.001

Robert Wood Johnson Foundation. (2006). Audit trail. Retrieved from http://www.qualres.org/HomeAudi-3700.html

Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice, 24*(2), 3-13. doi: 10.1111/j.1745-3992.2005.00006.x

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning taking memory tests improves long-term retention. Psychological Science, 17(3), 249-255. doi: 10.1111/j.1467-9280.2006.01693.x

Roedinger, H. L., & Marsh, E. J. (2005). The positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology, 31*(5), 1155-1159. Retrieved from http://psych.wustl.edu/memory/Roddy%20article%20PDF's/Roediger&Marsh2005.pdf

Rossiter, R. J. (2010, May 3). Fact Sheet: Nursing shortage. American Association of Colleges of Nursing. Retrieved from http://www.aacn.nche.edu

Rowland, L. (2005). *Test construction in an associate degree nursing program: A descriptive study*. (Master's thesis abstract). Retrieved from http://www.researchgate.net/publication

Sandahl, S. S. (2009). Collaborative testing as a learning strategy in nursing education. *Nursing Education Perspectives, 31*(3), 142-147. doi: 10.1043/1536-5026-31.3.142Scheele, T. H., Pruitt, R., Johnson, A., & Xu, Y. (2011). What do we know about educating Asian ESL nursing students? A literature review. *Nursing Education Perspectives, 32*(4), 244-249. doi:10.5480/1536-5026-32.4.244

Schroder, J. M. (2007). *A study of improving critical thinking skills with multiple-choice testing and first semester associate degree nursing students.* (Doctoral Dissertation). Retrieved from Proquest (3278154).

Seidman, I. (2006). *Interviewing as qualitative research* (3rd ed.). New York: Teachers College Press.

Shagrir, L. (2012). How evaluation processes affect the professional development of five teachers in higher education. *Journal of the Scholarship of Teaching and Learning, 12*(1), 23-35. Retrieved from iupui.edu/~josotl

Siegel, T. J. (2006). Do registered nurses perceive the anchoring of indwelling urinary catheters as a necessary aspect of nursing care?: A pilot study. *Journal of Wound Ostomy & Continence Nursing, 33*(2), 140-144.

Skelton, R. (2015). Quotations. Retrieved on March 8, 2015 from http://thinkexist.com/quotes/red_skelton/

Sonpal-Valias, N. (2009). Measuring the difference: An outcomes evaluation resources for the disability sector. Retrieved from htt://www.acd.ca

Stagnaro-Green, A. S., & Dowling, S. M. (2006). Used of flawed multiple-choice items by the New England Journal of Medicine for continuing medical education. *Medical Teacher, 28*(6), 566-568. doi:10.1080/01421590600711153

Stake, R. E. (1978). The case study method in social inquiry. *Educational Research, 7*(2), 5-8. Retrieved from http://links.jstor.org/sici?sici=0013-189X%28197802%297%3A2%3C5%3ATCSMIS%3E2.0.CO%3B2-I

Stake, R. E. (2005). Case studies. In N.K. Denzin & Y.S. Lincoln (Eds.) *Handbook of qualitative research* (2nd ed., pp 435-454). Thousand Oaks, CA: Sage

Stark, G. (2006). Stop "going over" exams! The multiple benefits of team exams. *Journal of Management Education, 30*(6), 818-827. doi: 10.1177/1052562906287965

Starr, K. (2009). Nursing education challenges: Students with English as an additional language. *Journal of Nursing Education, 48*(9), 478-487. doi: 10.3928/01484834-20090610-01

Starr, S. S. (2010). Associate degree nursing: entry into practice-link to the future. *Teaching and Learning in Nursing, 5*(3), 129-134. doi:10.1016/j.teln.2009.03.002

Stiggins, R. (2008). Assessment for learning, the achievement gap, and truly effective schools. Retrieved from http://www.murrieta.k12.ca.us/cms/lib5/CA01000508/Centricity/Domain/2241/A

ssessment%20for%20the%20learning%20gap%20in%20effective%20schools.pdf

Stiggins, R., & Du Four, R. (2009). Maximizing the power of formative assessment. *Phi Delta Kappan, 90* (9), 640-644. Retrieved from http://ehis.ebscohost.com.ezp.waldenulibrary.org/eds/pdfviewer/pdfviewer?vid=3&hid=120&sid=95f313f6-b956-4897-a3f6-cfdf3b9d9aad%40sessionmgr110

Stupans, I. (2006). Multiple-choice questions: Can they examine application of knowledge? *Pharmacy Education, 6*(1), 59-63. doi: 10.1080/1562210600567916.

Su, W. M. (2007). Writing content-dependent item sets that reflect critical thinking learning outcomes. *Nurse Educator, 32*(11), 11-15

Su, M. W., & Osisek, P. J. (2011). The revised Bloom's Taxonomy: Implications for educating nurses. *The Journal of Continuing Education in Nursing, 42*(7), 321-327. doi: 10.3928/00220124-20110621-05

Su, W. M., Osisek, P. J., Montgomery, C., & Peller, S. (2009). Developing multiple choice test items at higher cognitive levels. *Nurse Educator, 34*(5), 223-227.

Su, W. M., Osisek, P. J., & Starnes, B. (2005). Applying the revised Bloom's taxonomy in the clinical laboratory: Thinking skills involved in diagnostic reasoning. *Nurse Educator, 30* (3), 117-122. doi: 10.1097/00006223-200505000-00014

Su, W. M., Osisek, P. J., & Starnes, B. (2004).Applying the revised Bloom's taxonomy to a medical-surgical nursing lesson. *Nurse Educator, 29*(3), 116-120. doi: 10.1097/00006223-200405000-00010

Suskie, L.(2009). *Assessing student learning: A common sense guide* (2nd ed.). San Francisco: Jossey-Bass.

Tarrant, A. M., & Ware, J. (2011). Guidelines for developing high quality multiple-choice assessments. *Medical Education, 45*(3), 1-15 doi:10.111/j.1365-2923.2011 .04089.x

Tarrant, M., & Ware, J. (2010). A comparison of the psychometric properties of three- and four-option multiple-choice questions in nursing assessments. *Nurse Education Today, 30*(6), 539-543.doi.10.1016/j.nedt.2009.11.002

Tarrant, M., & Ware, J. (2008). Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. *Medical Education, 42*(2), 198-206. doi:10.1111/j.1365-2923.2007.02957

Tarrant, M., Knierim, A., Hayes, S., & Ware, J. (2006). The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. *Nurse Education Today, 26*(8). 662-671. doi:10.1016/j.nepr.2006.07.002

Tarrant, M., Ware, J., & Mohammed, A. M. (2009). An assessment of functioning and non-functioning distractors in multiple-choice questions: a descriptive analysis. *BMC medical education, 9*(1), 40. doi:10.1186/1472-6920-9-40

Taras, M. (2008). Summative and formative assessment: Perceptions and realities. *Active Learning in Higher Education, 9*(2), 172-192. doi: 10.1177/1469787408091655

Tiemeier, A. M., Stacy, Z. A., & Burke, J. M. (2011). Using multiple-choice questions at written at various Bloom's Taxonomy levels to evaluate student performance across a therapeutic sequence. *Innovations in Pharmacy, 2*(2), Article 41. Retrieved from http://z.umn.edu/INNOVATIONS

University of Kansas. (2013). What is a workshop? Retrieved from

http://ctb.ku.edu/en/tablecontents/sub_section_main_1113.aspx

Utley, R. (2011). *Theory and research for academic nurse educators*. Boston: Jones and Bartlett Publishers.

Vella, J. (1994). Learning to listen/learning to teach: Training trainers in the principles and practices of popular education. *Convergence, 27* (1), 5-22. Retrieved from http://ehis.ebscohost.com.ezp.waldenulibrary.org/

Vella, J. (2002). *Learning to listen, learning to teach: The power of dialogue in educating adults*. San Francisco: Jossey Bass

Vella, J. (2008). *On Teaching and Learning*. San Francisco CA: Jossey-Bass.

Waddell, D. L., & Dunn, N. (2005). Peer coaching: the next step in staff development. *Journal of Continuing Education in Nursing*, *36*(2), 84.

Walden Center for Research Support. Case Study Research Tutorial. Retrieved from http://researchcenter.waldenu.edu/Research_Tutorials/Case_Study_Research/index.html

Walker, J. (February 17, 2012). Brief introduction into metacognition. [You Tube video]. Retrieved from http://www.youtube.com/watch?v=mVE21QhY-lI

Walloch, J.A. (2006). *Assessment practices in the nursing classroom: An exploration of educators' assessment of students.* (Doctoral dissertation). Retrieved from Proquest. (3233950).

Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher*, 37(8), 469-479. doi: 10.3102/0013189X08327154

Webster-Wright, A. (2009). Reframing professional development through understanding authentic professional learning. *Review of Educational Research, 79*(2), 702-739. doi: 10.3102/0034654308330970

Weimer, M. (2012). How can I make my exams more about learning and less about grades? Magna 20 Minute Mentors.

Wendt, A., & Harmes, J.C. (2009). Evaluating innovative items for the NCLEX: Usability and pilot testing. *Nurse Educator, 3* (2), 56-59. Retrieved from https://m.ncsbn.org/Evaluating_Innovative_Items_pt1.pdf

Wendt, A., & Kenny, L. E. (2009). Alternative item types: Continuing the quest for authentic testing. *Journal of Nursing Education, 48*(3), 150-158. doi: 10.1097/NNE.0b013e31819fcae8

Wendt, A., & Kenny, L. E. (2007). Monitoring entry-level practice: Keeping the NCLEX –RN current. *Nurse Educator 32*(2), 78-80.

Wendt, A. (1999, Spring). Testing smarter with technology. *Popular Measuremen*t, 44-45. Retrieved from http://www.rasch.org/pm/pm2-44.pdf

Whitelock, D., & Cross, S. (2011). Academics understanding of authentic assessment. Retrieved from http://oro.open.ac.uk/29079/1/whitelocka-caa2011.pdf

Wiggins, G. (1989). A true test: Towards more authentic and equitable assessment. The *Phi Delta Kappan, 70*(9), 703-713. Retrieved from http://www.jstor.org/stable/20404004

Wiggins, G. (1989). *The case for authentic assessment*. *Practical Assessment, Research & Evaluation, 2*(2). Retrieved from http://eric.ed.gov

Wiggs, C. M. (2011). Collaborative testing: Assessing teamwork and critical thinking behaviors in baccalaureate nursing students. *Nurse Education Today*, 31(3), 279-282. doi: 10.1016/j.nedt.2010.10.027

Williams, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, *37*(1). 3-14. doi: 10.1016/stueduc.2011.03.001

Williams, J. B. (2006). Assertion-reason multiple-choice testing as a tool for deep learning: a qualitative analysis. *Assessment and Evaluation in Higher Education*, *31*(3), 287-301. doi: 10.1080/02602930500352857.

Willis, J. (2007). Assessment for learning-why the theory needs the practice. *International Journal of Pedagogies and Learning, 3*(2), 52-59. Retrieved from http://eprints.qut.edu.au/45078/1/Willis2007.pdf

Wo, A., & Gorham, J. L. (2009, Winter). Understanding the impact of enemy items on test validity and measurement precision, 15-17. *Clear Exam Review*, 14-18.

Wo, A., Wendt, A., & Lu, W. (2009, Winter). Readability of licensure examinations. *Clear Exam Review*, 21-23. Retrieved from https://www.ncsbn.org/Readability_of_Licensure_Examinations.pdf

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Washington D.C.: Sage.Yoho, M. J. (2010). Past and current trends of the NCLEX-RN. [Power Point slides]. Retrieved fromhttp://www.moadn.org/MOADN_Educational_Jan292010_ppts/Past%20and%20Current%20Trends%20in%20NCLEX-RN.ppt.

Appendix A: The Project

You are Invited to Attend a Faculty Professional Development Program

Who? ***All full time faculty***

What? ***Three day Workshop about Best Practices for the Assessment of Student Learning Outcomes***

Where: ***Skills Lab***

When: First week of faculty obligation

Why? ***Opportunity to network with your colleagues and create meaningful, authentic, and evidence-based assessments!***

Refreshments Provided….just bring an open mind!

Three day Faculty Professional Workshop followed by four Lunch and Learns **Best Practices for the Assessment of Student Learning Outcomes**

This workshop was designed using Vella's (2008) Seven Design Steps + One (How well).

Who?	Leader/Facilitator: Tracey Siegel MSN RN CNE Participants: 15 faculty members, and the Director of Nursing Education In 2 sessions (2 and 3) 3 faculty members will act as table leaders and facilitators.
Why?	Faculty have self-identified a need for development in the area of assessment of student learning outcomes.
When?	First week of faculty obligation during the fall or spring semester. This will be a three day workshop (21 hours) followed by four (4) 1 hour lunch and learn sessions for a total of 24 hours.
Where?	Main skills laboratory which is a smart classroom.
What?	Each session will be outlined below.
What for?	Achievement based program outcome: Following participation in this three day workshop and four monthly one hour lunch and learn sessions about best practices for the assessment of student learning outcomes, nursing faculty will demonstrate enhanced skills and competencies to design fair, reliable, and valid methods to assess student learning and assign grades. Individual session objectives are listed below.
How?	Learning tasks include an examination of current MCQ item analyses, creation of course specific test blueprint, revision of current MCQs, construction of new MCQ written at the applying and analyzing levels of Bloom's revised taxonomy and the formulation of a rubric for one course assignment. These will be explained in depth in the outline for each session.
How well? (Evaluation)	Following each day there will be a one hour reflective session to assess if the day's learning outcomes were met and to determine where the program needs to go during the next session. At the completion of the final workshop but before the start of the luncheon sessions, participants will be asked to complete a Professional Development Satisfaction Survey. This is mandated by the employing agency and has been modified. To determine the long term effectiveness of this program, I have planned a final lunch and learn which will be an opportunity for the faculty to honestly reflect upon the semester and determine if any of the new assessment strategies were incorporated into the curriculum. To establish if

	student learning outcomes have changed in relation to this program, student course-end evaluations will need to be assessed, analyzed, and trended each semester and linked to ongoing curriculum evaluation at the nursing program.

Day 1 Session One: *Understanding an item analysis and The How and Why of Test Blueprints*	Four hour session: 8am to 9:30 (1.5 hours + 1 hour spent viewing assigned webinar prior to attending session = 2.5 hours) Break 9:30 to 10 am 10am to 11:30 (1.5 hours) *Total time for session one: 4 hours* Lunch Break 11:30 to 12:30 (1 hour)
What?	1. 20 minute review of the process for item analysis using a PowerPoint presentation. 2. 20 minute overview of test blueprints- examples of three types used in nursing education using a PowerPoint presentation. 3. Remaining time spent working collaboratively designing test blueprints.
What for?	Following this session, faculty will have: 1. Interpreted a test item analysis for one MCQ examination. 2. Designed a test blueprint for one test used in one course. 3. Peer reviewed the test blueprints of the group.
How?	Learning Task One: Review the item analysis for one test. Use a highlighter and mark the questions that need revision (we will use these again in another session). Learning Task Two: Use one of the templates on the table to design a blueprint for one the tests in your course. Prior to this session, faculty will be requested to view *Test Blueprints: A Formula for Success* offered by *Nurse Tim* (nursing program subscribes to this service). This one hour webinar will provide faculty with information about how to design test blueprints. I will reinforce and introduce various types of blueprints in the PowerPoint presentation for faculty to appraise and decide which type works best for their individual courses. Following this, faculty will work in groups of 5 to create a workable blueprint for one test in each course. Materials: One each table faculty will have: Individual laptops with network capability Books: Mc Donald, M. E. (2013). The nurse educator's guide to assessing learning outcomes (3nd ed.). Boston, MA: Jones and Bartlett Publishers. Nilson, L. B. (2010). *Teaching at its Best.* San Francisco CA: Jossey Bass Oermann, M. H., & Gaberson, K. B. (2009). Evaluation and testing in nursing education (3rd ed.). New York: Springer Publishing Company.

	Suskie, L. (2009). Assessing student learning: A common sense guide (2nd ed.). San Francisco: Jossey-Bass. Articles: Notar, C. E., Zuelke, D. C., Wilson, J. D., & Yunker, B. D. (2004). The table of specifications: Insuring accountability in teacher made tests. Journal of Instructional Psychology, 31(2), 115-129. Other Resources: PowerPoint Presentation and handout from *Nurse Tim* presentation Copies of current test item analyzes, tests, test blueprints and all test materials from the four nursing courses. NCLEX-RN Test Plan (2013). Retrieved from https://www.ncsbn.org/2013_NCLEX_RN_TestPlan.pdf These resources will be available during the entire workshop. Suggested websites http://www.schreyerinstitute.psu.edu/Tools/ItemAnalysis/ http://www.washington.edu/oea/services/scanning_scoring/scoring/item_analysis.html http://www.jmu.edu/outreach/wm_library/Developing_Test_Blueprints.pdf http://www.uwlax.edu/catl/studentlearning/Presentations/Writing%20Better%20Objective%20Tests%20Handout.pdf

Session Two: *Good, Better, Best: Evaluating our own Multiple-Choice Questions*	Three hour session: 12:30pm to 2:00pm (1.5 hours) Break 2:00pm to 2:30 pm 2:30 pm to 4:00 (1.5 hours)
What?	Faculty will review, critique, and revise the highlighted questions from session one.
What for?	Following this session faculty will have: 1. Identified questions that do not meet best practices for item writing and test construction. 2. Revised at least 5 of these questions.
How?	Learning task: Use the rubric and resources on the tables to: 1. Review the questions you have highlighted in the first session. 2. Highlight the problem areas 3. Discuss these with your colleagues 4. Revise the questions based upon your discussions 5. Report your revisions to your colleagues at other tables. 6. Peer review revised questions.

	Materials: A holistic rubric (below) was designed to assist faculty to critique questions. Web-based resources and previously mentioned articles and books will be available. Suggested websites http://www.crlt.umich.edu/P8_0 http://cft.vanderbilt.edu/teaching-guides/assessment/writing-good-multiple-choice-test-questions/ Handout: NLN Board of Governors. (2012). The fair test imperative in nursing education. This is available at http://www.nln.org/aboutnln/livingdocuments/pdf/nlnvision_4.pdf

	Rubric
Good	Question clearly reflects 8 of the criteria as outlined by the National Council of State Boards of Nursing (NCSBN) as appropriate for submission to the NCLEX-RN Examination (*see definition of best*).
Better	Question clearly reflects between 9-11 of the criteria as outlined by the National Council of State Boards of Nursing (NCSBN) as appropriate for submission to the NCLEX-RN Examination (*see definition of best*).
Best	Question clearly reflects greater than 12 of the criteria as outlined by the National Council of State Boards of Nursing (NCSBN) as appropriate for submission to the NCLEX-RN Examination. Item is: 1. Focused on entry level nursing practice 2. Reflective of the legal scope of nursing practice 3. Current and relevant 4. Clear in language (no slang, idioms, brand names, or unfamiliar terms) 5. Editorially correct (spelling, grammar and punctuation). 6. Written at Bloom's revised applying and analyzing levels. 7. Linked to course objectives and program outcomes. Item has: 1. Only one correct response (key) and three plausible distractors. 2. Clearly stated intent and relevant, pertinent information. 3. Readability (can be read in 60 seconds and written at a 10^{th} grade level). 4. Consistent language. 5. A stem which poses a free standing question, prompts a positive response, and does not key students to correct answer. 6. Distractors which are plausible, and similar in length, detail or complexity; all

	options relate to a similar concept, and are arranged in a clear, logical order. 7. A key which is clear, concise, and the same length, detail, and complexity as the distractors. Item avoids: 1. Stereotyping, assumptions, ethnocentrism, elitism, or inflammatory material.

Session Three: Reflections on Day One	1 hour session (Total Hours Day 1=8 hours)
What?	Review what was accomplished today and outline what participants believe still needs to be accomplished.
What for?	Following this session faculty will have: 1. Described their feelings about today's sessions. 2. Discussed what needs to still be accomplished. 3. Planned for the next session.
How?	Learning Task One 1. Reflective Journal: Describe in writing what was accomplished today (15 minutes). 2. Discuss your findings with the larger group.

Day Two

Day 2: Session Three: *Abracadabra: A simple way to change multiple-choice questions to assess higher order thinking*	Four hour session: 8am to 9:30 Break 9:30 to 10 am 10am to 11:30 (1.5 hours) *Total time for session one: 4 hours* Lunch Break 11:30 to 12:30 (1 hour)
What?	Review of methods to convert remembering and understanding level MCQs to the applying and analyzing levels. The passing standard on the NCLEX-RN is applying and analyzing level questions. Many faculty have self-identified a knowledge deficit in this area.
What for?	Following this workshop faculty will have: 1. Differentiated questions written at various levels of Bloom's Revised Taxonomy. 2. Identified well written, plausible MCQ stems, keys, and distractors using the criteria

	outlined by McDonald (2013)
How?	Learning Task One: Use the resources on the table especially Mc Donald (2013) Appendices C, D, and E. 1. Using the MCQ that you have submitted for unit tests, select 5 that are written at the knowing and remembering levels and design stems that test at the applying and analyzing levels. 2. Discuss possible revisions with your peers. 3. Peer review the questions from colleagues at other tables. Articles Su, M. W., & Osisek, P. J. (2011). The revised Bloom's Taxonomy: Implications for educating nurses. The Journal of Continuing Education in Nursing, 42(7), 321-327. Su, W. M., Osisek, P. J., Montgomery, C., & Peller, S. (2009). Developing multiple choice test items at higher cognitive levels. Nurse Educator, 34(5), 223-227. Suggested Websites http://cft.vanderbilt.edu/teaching-guides/assessment/writing-good-multiple-choice-test-questions/
Session Four: *Test Reviews: Is There a Better Way?*	*3.5 hour session*: 12:30pm to 2:00pm (1.5 hours) Break 2:00pm to 2:30 pm 2:30 pm to 4:30 (2 hours)
What?	Introduction to other methods to conduct pre and posttest reviews. During the faculty focus groups it was identified by all participants that the current process "*doesn't work for the students or the faculty".*
What for?	Following this session faculty will have: 1. Differentiated among various methods to conduct pre and posttest reviews that are supported by the literature. 2. Discussed in small groups, strategies to incorporate these new methods into nursing courses.
How?	Learning Task One: 1. Watch 20 minute webinar *(How can I make my exams more about learning and less about grade*s?)

	2. Make a list of pro and cons of the methods presented 3. Decide how to incorporate one method into a course on a trial basis for the upcoming semester. Resources Walker, J. (February 17, 2012). Brief introduction into metacognition. [You Tube video]. Retrieved from http://www.youtube.com/watch?v=mVE21QhY-lI Suggested websites http://www.facultyfocus.com/articles/teaching-professor-blog/teaching-metacognition-to-improve-student-learning/
Session Five: Reflections of Day Two	1/2 hour session
What?	Review what was accomplished today and outline what participants believe still needs to be accomplished.
What for?	Following this session faculty will have: 1. Described their feelings about today's sessions. 2. Discussed what needs to still be accomplished. 3. Planned for the next session.
How?	Learning Task One 3. Think Pair Share: Discuss what was accomplished today with 2 other participants (10 minutes). 4. Discuss your findings with the larger group.

Day Three

Day 3 Session Five: *Assessment : Formative and Authentic Methods*	Two hour session: 8:00 to 10:00 Break 10:30 to 11:00
What?	Using a PowerPoint Presentation, I will review evidence-based authentic and formative methods to assess student learning outcomes and assign grades.
What for?	Following this session faculty will have: 1. Debated whether formative assessments can be seamlessly integrated into nursing courses based upon time constraints and the current

	format of the schedule. 2. Explained and supported their positions using the current evidence.
How?	Learning Task One: 1. Listen to 30 minute mini-lecture about formative and authentic assessment methods. Use the resources on the table to: 1. Debate the pro and cons of incorporating formative assessment methods into current course content. 2. Delineate one authentic assessment method currently used in the clinical component of each course that could be incorporated into the theory grade. Teaching Strategy: Debate Faculty will be randomly selected to debate the pros or cons of formative assessment. On the table there will be a folder with playing cards, based upon the cards selected faculty will be assigned one of the following position statements. Formative assessments are important and should be included into my course because: OR Incorporating formative assessments into course content would be difficult because: Handouts/Articles: Frey, B., & Schmitt, V. (2007). Coming to terms with classroom assessment. Journal of Advanced Academics, 18(3). Frey, B. B., Schmitt, V. L., & Allen, J. P. (2012). Defining authentic classroom assessment. Practical Assessment, Research, & Evaluation, 17(2). Stiggins, R. & Du Four, R. (2009). Maximizing the power of formative assessment. Phi Delta Kappan, 90(9), 640-644. Wiggins, G. (1989). A true test: Towards more authentic and equitable assessment. *The Phi Delta Kappan*, *70*(9), 703-713.

Session 6: Rubrics: *Building Blocks for Success*	1.5 hour session 11:00 to 12:30noon (1.5 hour) 12:30 to 1:30 Lunch
What?	Using a 20 minute PowerPoint I will provide an

	overview of the evidence supporting the use of rubrics in higher education .Nursing faculty have been slow to embrace these tools to grade or evaluate student assignments.
What for?	Following this session faculty will have: 1. Formulated one rubric (holistic or analytical) for one clinical nursing assignment.
How?	Learning Task One In Session 2 you had the opportunity to use a holistic rubric to evaluate MCQs. This tool gave you well defined guidelines to assist you as you revised questions. In your small group, select one clinical assignment, discuss what you expect from students completing that assignment, and define grading parameters. Use all of the resources to design one rubric (holistic or analytical) with at least three parameters. Then peer review the rubrics of your colleagues. Resources Suggested Websites: http://jfmueller.faculty.noctrl.edu/toolbox/rubrics.htm http://academic.pgcc.edu/~wpeirce/MCCCTR/Designingrubricsassessingthinking.html The Main Idea: Classroom Assessment for Student Learning: Doing it Right-Using it Well. Retrieved from www.TheMainIdea.net

Session 7: Reflections of Day 3	½ hour session 1:30 to 2pm
What?	Review what was accomplished today and what the participants feel still needs to be accomplished.
What for?	Following this session faculty will have: 1. Shared what they learned in their small groups. 2. Evaluated ongoing learning needs. 3. Planned for the three monthly lunch and learns.
How?	Learning Task One: On the provided index card complete the following sentence: After this two day professional development workshop I can now_____.

Lunch and Learn # 1 (October)

Journal Article Discussion	1 hour
What?	Discussion about the following article: Centrella-Nigro, A. M. (2012). Collaborative testing as posttest review. *Nursing Education Perspectives, 33*(5), 340-341.
What for?	Following this session faculty will have: 1. Summarized one article about the use of collaborative testing in nursing education as a posttest review. 2. Weighed the pros and cons of this strategy.
How?	Link to the article will be provided to faculty to facilitate reading prior to lunch. Learning Task One While you are reading this article, highlight important aspects that you would like to discuss with your colleagues.

Lunch and Learn # 2 (November)

Journal Article Discussion	1 hour
What?	Discussion about the following article: Bristol, T. J., & Secor, C. (2012). Clinical postconference online. *Teaching and Learning in Nursing, 7*(3), 123-126.
What for?	Following this session faculty will have: 1. Examined how one nursing program incorporated an online discussion board to promote critical thinking and clinical judgment. 2. Shared ideas about ways to integrate online discussion into a nursing course. 3. Reflected upon ways to use discussion boards as a formative assessment of learning.
How?	Link to the article will be provided to faculty to facilitate reading prior to lunch. Learning Task One While you are reading this article, highlight important aspects that you would like to discuss with your colleagues.

Lunch and Learn # 3 (1st week December)

Using Technology to Assist with the Assessment of Student Learning	1 hour webinar

What?	Webinar: Interactive Testing Technology from Nurse Tim. http://nursetim.com/webinars/Interactive_Testing_Technologies Webinar objectives: • Describe basic principles of assessment strategies with technology. • Utilize two technologies that can be used for assessment of student competency. • Develop policy that will facilitate successful management of testing technologies.
What for?	Following this session faculty will have: 1. Viewed an online presentation about ways to integrate Blackboard and other online course management systems to provide alternative methods for the assessment of student learning outcomes. Included in this webinar are tips for making in class paper and pencil tests and online tests more reflective of the NCLEX-RN content.
How?	All faculty already have access to this webinar. Viewing it together will allow a discussion about how technology may assist to make faculty-developed examination more reliable, valid, and reflective of NCLEX –RN content and methods.

Lunch and Learn # 4 (End of December at conclusion of semester)

Reflections and Evaluation of Faculty Development Workshops and Luncheon Sessions	1 hour
What?	Evaluation of this faculty professional development session using reflective questions.
What for?	Following this session faculty will have: 1. Identified how new assessment methods were integrated into courses. 2. Summarized the results of changes to MCQ examinations. 3. Reflected upon ongoing faculty development needs.
How?	Learning Task One: Please respond to the following questions. 1. What new assessment practice did you integrate into your course this semester? 2. How did it go?

	3. Will you do it again next semester? 4. What was the student feedback? 5. What will you need to revise for next time? 6. What do you still want to know about student assessment?

References

Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, *53*(1), 27-31.

Chappuis, S., Chappuis, J., & Stiggins, R. (2009). The Quest for Quality. *Educational Leadership*, *67*(3), 14-19.

Daggett, L. M. (2008). A rubric for grading or editing student papers. *Nurse Educator*, *33*(2), 55-56.

Oermann, M. H., & Gaberson, K. B. (2009). Evaluation and testing in nursing education (3rd ed.). New York: Springer Publishing Company.

Penn, B. (February 8, 2013). Developing and analyzing test items. Retrieved from http://www.aacn.nche.edu/membership/members-only/presentations/2013/13facdev/Penn.pdf

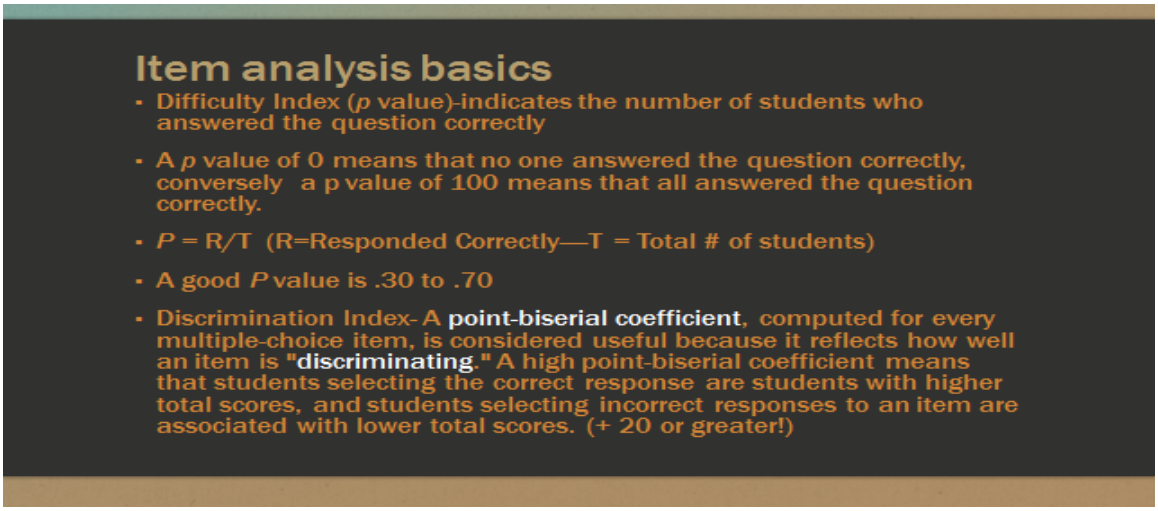
Quigly, L. (2012). May College Workshop: Authentic Assessment. Retrieved from www.sage.edu/centers/asset/Authentic_Assessment.pd

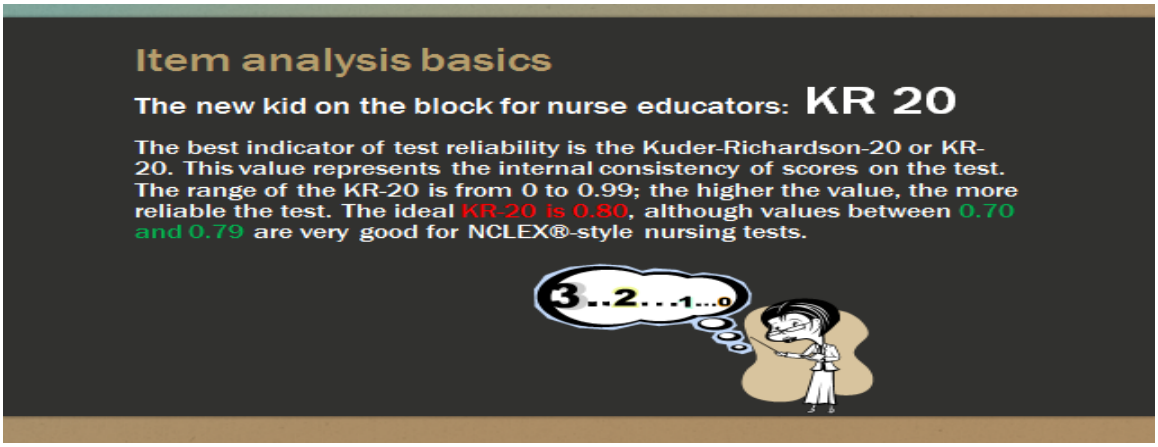
Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment and Evaluation in Higher Education*, *35*(4), 435-448.

Stevens, D. D, & Levi, Antonia J. (2005). *An Introduction to Rubrics*. Sterling VA: Stylus

Presentation 1 Session 1

Item Analysis and Blueprints





Remember!

- Mastery Questions: It is OK to have 100 % of students answer a question correctly if that question is something that needs to be mastered.
- Don't expect a normal bell curve on teacher made tests (Oermann & Gaberson, 2009) especially in nursing programs- we have highly qualified students who are "above average"
- Nurse educators typically use criterion-referenced scoring (we compare to a score of 100-not against fellow students)

Test Blueprinting Essentials (Oermann and Gaberson, 2009; p. 66) Also called test specifications!

1. A list of major topics or instructional objectives the test will cover
2. The level of complexity of the task to be assessed
3. The emphasis each topic will have indicated by the number or % of items or points

Content validity is ensured by closely following the test blueprint which reflects well-written objectives

You can design using objective or topics

Is this why some of our student think there is a disconnect between what we teach and what is on the test?

http://vimeo.com/68331396

Suggestions from Penn (2013) and the AACN

- ✓ Draft overview of course objectives, units, or content (in collaboration with colleagues)
- ✓ Estimate relative teaching time
- ✓ Transfer these proportions to the test plan
- ✓ Consider all steps in the nursing process
- ✓ Include all possible levels of the cognitive domain

Sample Test Blueprint Antihypertensive Medications Two way table 40 question test

Content	Knowing	Understanding	Applying	Totals
Beta Blockers (25%)	2	4	4	10
Diuretics (25%)	2	4	4	10
ACE/ARBs (25%)	2	4	4	10
Ca Channel Blockers (12.5%)	1	2	2	5
Other (12.5%)	1	2	2	5
Totals (100%)	8	16	16	40

Three way table of specifications: same test

Outcomes/Process	Remember (20%)	Understand (40%)	Apply (40%)	Total
Assessment (40%)	D, BB x2, ACE, O, CA,	O x 2, BB, ACE, ARB, D	CA, O, D, BB	16
Diagnosis (10%)	O	D, BB, CA		4
Planning (10%)	D	D, CA	O	4
Interventions (20%)		BB, ACE	ACE, ARB, D x 2 O, BB	8
Evaluation (20%)		D, O, ACE	O x 2, BB, CA, ACE	8
	8	16	16	40

Key: BB=beta blocker, ACE, ARB, CA = calcium channel blocker, D=diuretics, O=other

Now its your turn

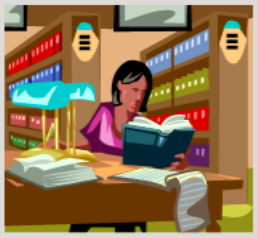
Questions?

Day 3 Session 5

Authentic and Formative Assessments

FORMATIVE AND AUTHENTIC ASSESSMENTS

Faculty Development Workshop
Tracey Siegel MSN RN CNE



What do we mean by authentic assessment?

- "Assessment is authentic when we directly examine student performance on worthy, intellectual tasks" (Wiggins, 1990).
- An assessment that mirrors or reflects real world applications or the performance of tasks in realistic settings.
- In nursing we have the ultimate "authentic assessment" and we call it clinical!

Examples of Authentic Assessments in Nursing Education

Our students:

- Construct Care Plans
- Interpret interpersonal process recordings
- Design grand rounds case studies
- Perform physical assessments
- Document nursing care
- Demonstrate nursing skills and clinical judgment during clinical and clinical simulations!

- Yet....we do not assign a grade to any of these meaningful and real world tasks!
- More about this later!

Formative Assessment

- A process that teachers use to provide feedback to students and determine students' ongoing learning needs.
- Stiggins (2008) has termed this process assessment for learning rather than the traditional approach which is the assessment of learning.
- Suskie (2009) calls it those assessments that are undertaken while student learning is taking place, rather than at the end of a course.
- Research has confirmed that formative assessment is often not a well-established practice in many nursing education programs (Duers & Brown, 2009; Koh, 2010; Mayya, 2010; Miller, 2012).

So are our current MCQ examinations formative?

- Should we incorporate other methods into our courses to allow students the opportunity to make mistakes and learn from then in a low stakes way?
- Would these assessments also help us determine what the students did and did not learn without a high stakes test?

Examples of formative assessment methods used in nursing education

- Quizzes (paper and pencil or online)
- Online Discussion Boards
- Blogs
- Games
- Concepts Maps
- Case Studies
- Role Play
- Debates
- Reflective Journals
- Clickers
- E- Portfolios
- Socratic Questioning
- CATs (Classroom Assessment Techniques)

- Formative assessments may or may not be graded!
- It depends upon the method and reasons!
- But all should tell us where the student is on the learning continuum!

In your assigned groups debate the following and come back in 30 minutes and report your findings to all of us!

Resolved:

Formative assessments are important and should be included into my course because:

Resolved:

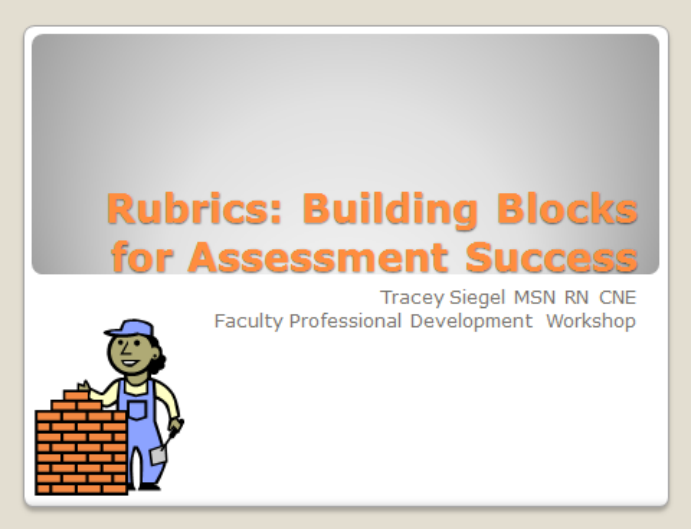
Incorporating formative assessments into course content while important would be difficult because:

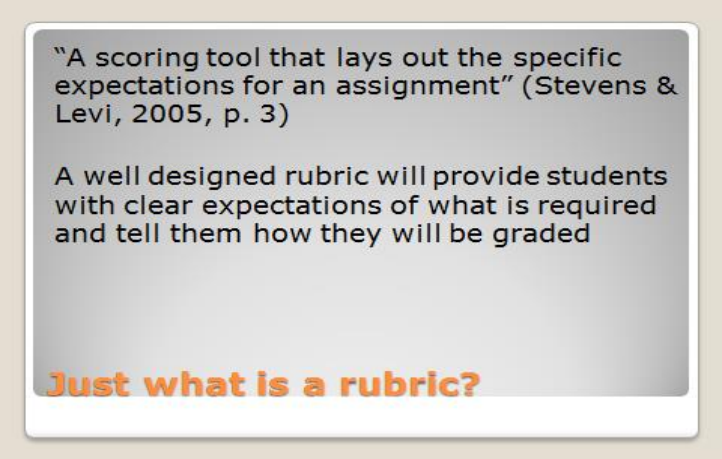
SO…WHERE DO WE GO FROM HERE?

Rubrics!

Day 3: Session 6

Rubrics





Examples of assignments:

- Reflective journals
- Online discussions
- Research papers
- Group work
- Oral presentations
- Almost any assignment that needs clear guidelines for both you and your students!

When should I use one?

According to Stevens and Levi (2005) you need them if:

- ✓ You are getting writers cramp from writing the same comments over and over
- ✓ It takes you an entire class to explain your expectations for an assignment
- ✓ You fear that you are not being objective…the first papers are critically appraised and by the last you cannot even remember what you expected!

Why would I need one?

A rubric may help with this!

Checklists are the simplest to create and use but focus on presence rather than quality. They consist of a 2 column table with the component on the left and then room for check mark on the right.

We have used basic ones in the past!

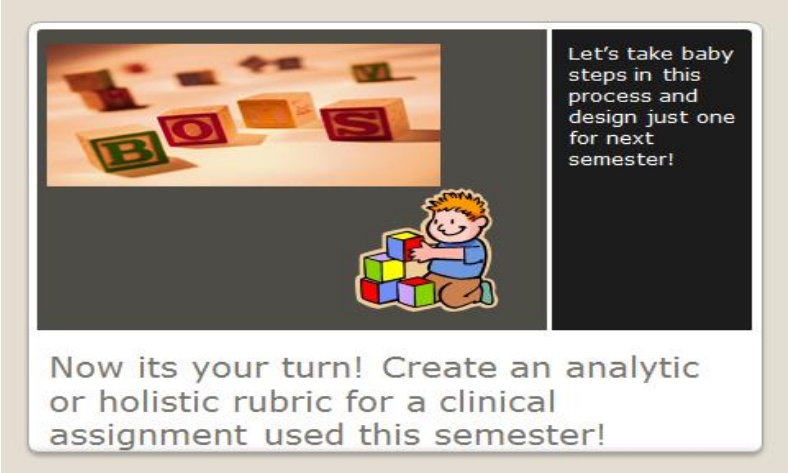
Criteria	4 points	3 points	2 points	1 point
Has a plan for Investigation	The plan is thorough	The plan is lacking a few details	The plan is missing major details	The plan is incomplete and limited
Use of Materials	Manages all materials responsibly	Uses the materials responsibly most of the time	Mishandles some of the materials	Does not use materials properly
Collects the Data	Thorough collection	Some of the data	Major portions of the data are missing	The data collection consists of a few points

One example of analytic rubric- more in your online resources


Preview Rubric
Demo Holistic Rubric

Level	Description	Feedback
Superior 75 % or more	• Provides ample supporting detail to support solution/ argument. • Organizational pattern is logical and conveys completeness. • Uses effective language; makes engaging, appropriate word choices for audience and purpose. • Consistently follows the rules of standard English.	Great job.
Excellent 50 % or more	• Provides adequate supporting detail to support solution/argument. • Organizational pattern is logical & conveys completeness & wholeness with few lapses. • Uses effective language & appropriate word choices for intended audience & purpose. • Generally follows the rules for standard English.	Good job.
Adequate 25 % or more	• Includes some details, but may include extraneous or loosely related material. • Achieves little completeness & wholeness though organization attempted. • Limited & predictable vocabulary, perhaps not appropriate for intended audience & purpose. • Generally does not follow the rules of standard English.	Needs improvement, but good effort.
Inadequate 0 % or more	• Includes inconsistent or few details which may interfere with the meaning of the text. Little evidence of organization or any sense of wholeness & completeness. • Has a limited or inappropriate vocabulary for the intended audience & purpose. • Does not follow the rules of standard English.	Needs to be re-written.

Holistic Rubric- you used one in session 2! More in your online resources


Let's take baby steps in this process and design just one for next semester!
Now its your turn! Create an analytic or holistic rubric for a clinical assignment used this semester!

Faculty Professional Development Satisfaction Survey: Assessment Workshop

Please complete and return to facilitator at the conclusion of this workshop.

Overall Evaluation	**Excellent**	**Good**	**Fair**	**Poor**
Workshop Quality				
Workshop Content				
Relevance to my job				
Usefulness to my job.				
Workshop Outcomes	**Strongly Agree**	**Agree**	**Disagree**	**Strongly Disagree**
I learned useful strategies.				
I plan to implement the strategies I have learned next semester.				
Presenter aligned the objectives of this workshop with my professional needs.				
Presenter prepared me to implement new strategies.				
Materials were pertinent and useful.				
A supportive environment of a professional learning community was created.				
Opportunities to network and learn from colleagues were supported.				

What follow up assistance do you need in the area of assessing student learning outcomes?

Do you think this workshop will ultimately affect student learning outcomes?

Additional Comments:

Appendix B: Letter of Cooperation and Data Use Agreement

January 2012

Dear Tracey Siegel,

Based on my review of your research proposal, I give you permission to conduct the study entitled Assessment Practices at an Associate Degree Nursing Program: Study within the XXXX Nursing Program As part of this study, I authorize you to:

Interview program administration and faculty using individual interviews or focus groups,

Review a sample of teacher-developed multiple choice examinations,

Review test blueprints and completed test item analyses for the sample of teacher-developed multiple choice examinations,

Analyze course-end evaluations completed anonymously online by students,

Request (2) course coordinators complete reflective diaries during item writing and test construction.

This information must be kept confidential and shared only with your project chairpersons. I understand that this research study will be published as a capstone project and possibly in a nursing education peer reviewed journal at some future date. Individual faculty participation will be voluntary and at their own discretion.

I understand that XXXX nursing program responsibilities include allowing you to review various documents as outlined above. Additionally, I am allowing you the use of the faculty conference room or skills laboratory to conduct the faculty focus groups, and use of administrators' offices to conduct individual face-to-face interviews. We ask that

faculty focus groups be conducted after regularly scheduled classroom and clinical responsibilities. XXXX Nursing Program reserves the right to withdraw from this study at any time if circumstances change or need arises.

I confirm that I am authorized to approve research in this setting. I understand that the data collected will remain entirely confidential and may not be provided to anyone outside of the research team without permission from the Walden University IRB.

Sincerely,

Authorization Official

Contact Information

Appendix C: Consent Forms

Course Coordinator Diaries, Interview and Focus Groups

Diary Consent

Title of the Study

Assessment Practices at an Associate Degree Nursing Program:

A Case Study

You are asked to participate in a qualitative research case study conducted by Tracey Siegel, a fellow faculty member, as part of a doctoral capstone project for Walden University. Your participation in this study is entirely voluntary. Please read the information below and ask questions about anything you do not understand before deciding whether or not to participate. This form is part of a process called "informed consent" to allow you to understand this study before deciding whether to take part. You may already know the researcher as a fellow faculty member and colleague however, this study is separate from that role and your participation will not impact your employment in any way.

You have been asked to participate because you are a member of the faculty of the XXXX Nursing Program and are involved with multiple-choice question item writing and test construction. You are also responsible to assess and evaluate student learning outcomes.

Purpose of the Study

The purpose of this study is to describe the faculty and student issues related to using teacher-developed MCQ examinations as the only method to assess students' knowledge of nursing theory and to assign grades. The following question will guide this research: What are the issues related to the use of teacher-developed MCQ examinations as the only method of student assessment and evaluation in the theory component of courses at a suburban associate degree nursing program? Little is known about how this assessment method affects students, influences the practice of the faculty, or impacts the learning outcomes of the nursing program.

Procedures

If you volunteer to participate in this study, you will be asked to maintain a diary or reflective journal to describing feelings prior to, during, and following each examination

over the course of one semester (**Fall 2012**). You will also include any issues or problems related to these examinations. These will be collected and analyzed by the researcher for common themes or concepts. After your responses are transcribed by the researcher you will be asked by the research to verify if your statements (member check) are an accurate reflection of your opinions or perceptions. If not, you will have the opportunity to make additional statements or corrections.

Voluntary Nature of the Study

This study is voluntary. The researcher will respect your decision of whether or not you choose to be in the study. No one at XXXX Nursing Program or the Medical Center will treat you differently if you decide not to be in the study. If you decide to join the study now, you can still change your mind during or after the study. If you volunteer to be in this study, you may withdraw at any time without consequences of any kind or loss of benefits to which you are otherwise entitled.

Risks and Benefits of Being in the Study

Maintaining a diary may impact your professional or personal time. Being in this study would not pose risk to your safety, wellbeing, or job security.

Participating in this research will assist this researcher to understand the teacher and student issues related to using only teacher-developed MCQ examinations. This may ultimately help you as faculty to determine the best teaching and learning practices for the XXXX Nursing Program.

Payment for Participation

Course Coordinators who agree to maintain a reflective journal will be given a $5.00 College Bookstore Gift Card as a thank you for the time involved.

Privacy

Any information you provide will be kept confidential. During the coding and data analysis process, all participants will be identified by a pseudonym. The researcher will not use your personal information for any purposes outside of this research project. Also, the researcher will not include your name or anything else that could identify you in the study reports. All data will be stored securely in a locked file cabinet located at the researcher's home and kept for a period of at least 5 years as required by Walden University.

Contacts and Questions

You may ask any questions you have now. Or if you have questions later, you may contact the researcher via cell phone (xxx-xxx-xxxx) or email (xxxx.xxxx@xxx.xxx). address. If you want to talk privately about your rights as a participant, you can call Dr. Leilani Endicott. She is the Walden University representative who can discuss this with you. Her phone number is 1-800-xxx-xxxx, extension xxxx.

The researcher will give you a copy of this form to keep.

Statement of Consent

I have read the above information and I feel I understand the study well enough to make a decision about my involvement. By signing this consent form I understand that I am agreeing to the terms described above.

Printed Name of Participant: Date of consent:

Participant's Signature: Researcher's Signature:

Individual Interview

Title of the Study

Assessment Practices at an Associate Degree Nursing Program:

A Case Study

You are asked to participate in a qualitative research case study conducted by Tracey Siegel, a fellow faculty member, as part of a doctoral capstone project for Walden University. Your participation in this study is entirely voluntary. Please read the information below and ask questions about anything you do not understand before deciding whether or not to participate. This form is part of a process called "informed consent" to allow you to understand this study before deciding whether to take part. You may already know the researcher as a fellow faculty member and colleague however, this study is separate from that role and your participation will not impact your employment in any way.

You have been asked to participate because you are a member administration of the XXXX Nursing Program and are involved with multiple-choice question item writing and test construction. You are also responsible to assess and evaluate student learning outcomes.

Purpose of the Study

The purpose of this study is to describe the faculty and student issues related to using teacher-developed MCQ examinations as the only method to assess students' knowledge of nursing theory and to assign grades. The following question will guide this research: What are the issues related to the use of teacher-developed MCQ examinations as the only method of student assessment and evaluation in the theory component of courses at a suburban associate degree nursing program? Little is known about how this assessment method affects students, influences the practice of the faculty, or impacts the learning outcomes of the nursing program.

Procedures

If you volunteer to participate in this study, you will be asked to participate in a single individual one hour audiotaped face-to-face interview with the researcher to ascertain your opinions of the current assessment practices at your program.

After your responses are transcribed by the researcher you will be asked by the research to verify your statements (member check) to make sure they are an accurate reflection of

your opinions or perceptions. If not, you will have the opportunity to make additional statements or corrections.

Voluntary Nature of the Study

This study is voluntary. The researcher will respect your decision of whether or not you choose to be in the study. No one at XXXX Nursing Program or the Medical Center will treat you differently if you decide not to be in the study. If you decide to join the study now, you can still change your mind during or after the study. If you volunteer to be in this study, you may withdraw at any time without consequences of any kind or loss of benefits to which you are otherwise entitled. You may also refuse to answer any questions you do not want to answer during the interviews or focus groups. There is no penalty if you withdraw from the study and you will not lose any benefits to which you are otherwise entitled.

Risks and Benefits of Being in the Study

Being in this type of research study involves the remote possibility of becoming upset or anxious during the interview; therefore you can choose not to answer any question that you find upsetting. This study may also impact your personal time as the interview will be conducted after scheduled class time as to not affect the normal schedule and routines at XXXX College. Being in this study would not pose risk to your safety, wellbeing, or job security. Participating in this research will assist this researcher to understand the teacher and student issues related to using only teacher-developed MCQ examinations. This may ultimately help you to determine the best teaching and learning practices for the XXXX College Nursing Program.

Payment for Participation

There is no payment provided for participation in this interview.

Privacy

Any information you provide will be kept confidential. During the coding and data analysis process, all participants will be identified by a pseudonym. The researcher will not use your personal information for any purposes outside of this research project. Also, the researcher will not include your name or anything else that could identify you in the study reports. All data will be stored securely in a locked file cabinet located at the researcher's home and kept for a period of at least 5 years as required by Walden University. The digital audiotapes of the interview will remain on my home computer and on a flash drive for 5 years, and then deleted. I ask that you also respect the privacy and

anonymity of other research participants and not divulge or reveal the contents of the interview.

Contacts and Questions

You may ask any questions you have now. Or if you have questions later, you may contact the researcher via cell phone (xxx-xxx-xxxx) or email (xxxx.xxxx@xxx.xxx). address. If you want to talk privately about your rights as a participant, you can call Dr. Leilani Endicott. She is the Walden University representative who can discuss this with you. Her phone number is 1-800-xxx-xxxx, extension xxxx. Walden University's approval number for this study is 08-02-12-0156784 and it expires on August 1, 2013.

The researcher will give you a copy of this form to keep.

Statement of Consent

I have read the above information and I feel I understand the study well enough to make a decision about my involvement. By signing this consent form I understand that I am agreeing to the terms described above.

Printed Name of Participant: Date of consent:

Participant's Signature: Researcher's Signature:

Focus Groups

Title of the Study

Assessment Practices at an Associate Degree Nursing Program:

A Case Study

You are asked to participate in a qualitative research case study conducted by Tracey Siegel as part of a doctoral capstone project for Walden University. Your participation in this study is entirely voluntary. Please read the information below and ask questions about anything you do not understand before deciding whether or not to participate. This form is part of a process called "informed consent" to allow you to understand this study before deciding whether to take part. You may already know the researcher as a fellow faculty member and colleague however, this study is separate from that role and your participation will not impact your employment in any way.

You have been asked to participate because you are a member of the faculty or administration of the XXXX College Nursing Program and are involved with multiple-choice question item writing and test construction. You are also responsible to assess and evaluate student learning outcomes.

Purpose of the Study

The purpose of this study is to describe the faculty and student issues related to using teacher-developed MCQ examinations as the only method to assess students' knowledge of nursing theory and to assign grades. The following question will guide this research: What are the issues related to the use of teacher-developed MCQ examinations as the only method of student assessment and evaluation in the theory component of courses at a suburban associate degree nursing program? Little is known about how this assessment method affects students, influences the practice of the faculty, or impacts the learning outcomes of the nursing program.

Procedures

If you volunteer to participate in this study, you will be asked to participate in a single one hour focus group conducted by the researcher. During this focus group you will be asked to discuss your perceptions of teacher-developed multiple-choice question tests and other methods of formative and summative student assessment. The number of participants in each focus group will depend upon the number of faculty who volunteer to participate. I anticipate that each focus group will have 6-7 participants. This focus

group will be audiotaped with a digital recorder and notes will be taken by the researcher. After the responses are transcribed by the researcher you will be asked to read the transcript to make sure that your statements are an accurate reflection of your opinions or perceptions. If not, you will have the opportunity to make additional statements or corrections.

Voluntary Nature of the Study

This study is voluntary. The researcher will respect your decision of whether or not you choose to be in the study. No one at the Nursing Program or Medical Center will treat you differently if you decide not to be in the study. If you decide to join the study now, you can still change your mind during or after the study. If you volunteer to be in this study, you may withdraw at any time without consequences of any kind or loss of benefits to which you are otherwise entitled. You may also refuse to answer any questions you do not want to answer during the interviews or focus groups. There is no penalty if you withdraw from the study and you will not lose any benefits to which you are otherwise entitled.

Risks and Benefits of Being in the Study

Being in this type of research study involves the remote possibility of becoming upset or anxious during the focus group or interview; therefore you can choose not to answer any question that you find upsetting. This study may also impact your personal time as the focus groups will be conducted after scheduled class time as to not affect the normal schedule and routines at the College. Being in this study would not pose risk to your safety, wellbeing, or job security.

Participating in this research will assist this researcher to understand the teacher and student issues related to using only teacher-developed MCQ examinations. This may ultimately help you as faculty to determine the best teaching and learning practices for the XXXX College Nursing Program.

Payment for Participation

Since the focus group will be after scheduled classes and may occur during the dinner hour, refreshments such as sandwiches or pizza will be provided by the researcher.

Privacy

Any information you provide will be kept confidential. During the coding and data analysis process, all participants will be identified by a pseudonym. The researcher will

not use your personal information for any purposes outside of this research project. Also, the researcher will not include your name or anything else that could identify you in the study reports. All data will be stored securely in a locked file cabinet located at the researcher's home and kept for a period of at least 5 years as required by Walden University. The digital audiotapes of the focus groups and interviews will remain on the researcher's home computer and on a flash drive for 5 years, and then deleted. I ask that you also respect the privacy and anonymity of other research participants and not divulge or reveal the contents of the interviews or focus groups.

Contacts and Questions

You may ask any questions you have now. Or if you have questions later, you may contact the researcher via cell phone (xxx-xxx-xxxx) or email (xxx.xxx@xxx.xxx). address. If you want to talk privately about your rights as a participant, you can call Dr. Leilani Endicott. She is the Walden University representative who can discuss this with you. Her phone number is 1-800-xxx-xxxx, extension xxxx. The researcher will give you a copy of this form to keep.

Statement of Consent

I have read the above information and I feel I understand the study well enough to make a decision about my involvement. By signing this consent form I understand that I am agreeing to the terms described above.

Printed Name of Participant: Date of consent:

Participant's Signature: Researcher's Signature:

.

.

Appendix D: Case Study Protocol

Purpose of the Protocol

The purpose of this protocol is to outline the procedures of this case study. This will help to provide clear guidelines and to ensure the reliability of this research.

Guiding Question

What are the issues related to the use of teacher-developed MCQ examinations as the only method of student assessment and evaluation in the theory component of courses at a suburban associate degree nursing program?

Sub questions:

1. How has using only teacher-developed MCQ examinations impacted the teaching/learning process?
2. How has using only teacher-developed MCQ examinations impacted the teacher/student relationship?
3. How are students assessed formatively in the theory component of the nursing courses?

Theoretical Frameworks

1. National League for Nursing's Core Competencies of Nursing Educators
2. Anderson's and Krathwohl's revised Blooms taxonomy

Data Collection Procedures

Site

The setting for this research is an associate degree nursing program located in an eastern Mid-Atlantic State within the United States of America. This program is owned by a community college but staffed by nursing faculty who are employees of a local nonprofit community hospital.

Possible Participants

All of the faculty are women between the ages of 40 and 65. Only three faculty members are ethnically and culturally diverse. One is from India and has lived in the United States for 25 years; one is African American and has lived in the United States

since birth and the other is from Cuba and has lived in the United States for 35 years. The faculty member from Cuba did not participate.

Participants	**Educational Preparation/Credentials**	**Job Description**	**Years as Academic Nurse Educators**
Administrator Number 1 Nadine	EdD credits, MSN, BSN Certified Nurse Educator	Director of Nursing Education	27
Danielle	MSN, BSN Certified Nurse Educator	Course Coordinator	20
Diane	MSN, BSN	Course Coordinator	22
Sharon	MSN, BSN Certified Nurse Educator	Course Coordinator	11
Coleen	MSN, BSN	Course Coordinator	35
Megan	MSN, BSN	Instructor	4
Sandra	MSN, BSN	Instructor	8
Beth	MSN, BSN	Instructor	8
Jane	MSN, BSN	Instructor	20
Laura	MSN, BSN	Instructor	1

Data Collection Plan (multiple sources of evidence)

1. Interview with nursing program administration,
2. Focus groups with faculty members,
3. Document review (test blueprints, sample of test questions, test item analyses, course end evaluations),
4. Personal diaries of 2 course coordinators involved with test construction maintained during a semester or immediately following the semester.

Data Analysis

1. Open coding,
2. Categorical analysis

Appendix E: Course Coordinator Journal Protocol

As the course coordinator you will be involved with test construction during this semester. Please answer the following questions prior to and after the administration of each scheduled examination and the final. You can answer these questions directly on this form or use whatever method is most convenient for you.

Questions before each scheduled test:

1. How do you feel about the quality of the multiple-choice questions (MCQ) included on this test?
2. Did you have difficulty writing or finding well-written MCQ for this test?
3. Is there a blueprint for all MCQ which corresponds to course objectives and accurately reflects Bloom's Taxonomy and the NCLEX-RN Test plan?

Questions after each scheduled test:

1. Were there any issues during test administration related to the MCQ (for example, were there problems with form and style or issues with student comprehension of questions during the test)?
2. After item analysis, did you need to change the key (accept two answers) or "throw out" questions?
3. Were there any other issues related to this test?
4. How valuable was the "test review" for students?
5. Were there any issues related to the test that came to light during the "test review"?

Appendix F: Interview Protocol

Participant: The Director of Nursing Education

1. Introduce self
2. Describe project and goals
3. Have participant sign consent.

The interviews will be audiotaped. It will be semistructured and the questions listed below are representative of the ones I will ask, but I expect other questions to emerge during the course of the interviews. I also plan to use detail oriented elaboration and clarification probes or prompts. Examples include:

1. Please give me a specific example.
2. Please tell me more about that.
3. I'm not sure I understand what you mean by "x". Would you explain that further?

Opening Statement

I want to thank you for agreeing to meet with me today. As you know, I am finishing my doctorate at Walden University and as part of my capstone project I am researching the issues related to using teacher-developed multiple-choice question examinations as the only method of student assessment and grading in the theory component of courses at this nursing program. Nursing education has traditionally used this method, however in recent years some educational leaders have called for new and more innovative ways to assess student learning outcomes. However, others believe that multiple choice question tests are the best way to prepare students for the licensure examination and feel that it is adequate in most situations. I want to explore with you

your opinions and feelings about the teacher-developed multiple-choice question exams used at this program. This interview will take approximately one hour. I need to audiotape so I do not miss any of your comments; all responses and comments will be kept confidential. When I write the narrative for my research, you will be identified by a pseudonym. Remember, you do not have to talk about anything you do not want to, but I encourage you to be open and honest as I value your opinions and perceptions.

The information you provide will be used to help me understand this issue and may possibly help all of us as we try to use evidence-based assessment and evaluation methods in this nursing program. A transcript of this interview will be shared with both of you. At that time you will be asked to reflect upon your responses and make any changes or clarifications to ensure that I fully understand your comments.

Part I

Introductory Questions

1. What was the impetus for using the teacher-developed multiple choice question examinations as the only method to assess students in the theory component of nursing courses?
2. Do you believe that this method adequately reflects the student learning outcomes?
3. Please tell me about your experiences with the teacher-developed MCQ examinations used at the college.

Part II

Exploring the impact on the teaching/learning process?

1. How has using one method to assess students and assign grades impact the way the faculty teaches?
2. How has using one method to assess students and assign grades impact the students?
3. How has using this method impacted student retention?
4. How has using this method impacted students with limited English proficiency and those who were educated in foreign schools?

Part III

Exploring the role of assessment

1. How are students assessed formatively in the theory component of nursing courses?
2. In your opinion, how comfortable are faculty with using other methods of assessing student learning?

Appendix G: Focus Group Protocol

Participants: Faculty of the Nursing Program. Depending upon how many agree to participate, each focus group will contain six or seven participants. I will serve as the moderator.

1. Introduce self
1. Describe project and goals
2. Have participants sign consents.

Focus groups will be audiotaped. They will be semistructured and the questions listed below are representative of the ones I will ask, but I expect other questions to emerge during the course of the interviews. I also plan to use detail oriented, elaboration, and clarification probes or prompts. Examples include:

1. Please give me a specific example.
2. Please tell me more about that.
3. I'm not sure I understand what you mean by "x". Would you explain that further?

Opening Statement

I want to thank you for agreeing to meet with me today. As you know, I am finishing my doctorate at Walden University and as part of my capstone project I am researching the issues related to using teacher-developed multiple-choice question examinations as the only method of student assessment and grading in the theory component of courses at this nursing program. Nursing education has traditionally used this method, however in recent years some educational leaders have called for new and more innovative ways to assess student learning outcomes. However, others believe that

multiple choice question tests are effective in prepare students for the licensure examination and feel that it is adequate in most situations.

I want to explore with all of you your opinions and feelings about the teacher-developed multiple-choice question exams used at this program. This focus group will take approximately one hour. I need to audiotape so I do not miss any of your comments; all responses and comments will be kept confidential. When I write the narrative for my research, you will be identified by a pseudonym. Remember, you do not have to talk about anything you do not want to, but I encourage you to be open and honest as I value your opinions and perceptions. A transcript of the focus group will be shared with all of you. At that time you will be asked to reflect upon your responses and make any changes or clarifications to ensure that I fully understand your comments.

The information you provide will be used to help me understand this issue and may possibly help all of us as we try to use evidence-based assessment and evaluation methods in this nursing program.

Research Question

What are the issues related to the use of teacher-developed multiple-choice question examinations as the only method of student assessment and evaluation in the theory component of courses at XXXX College?

Additional research questions:

1. How has using only teacher-developed MCQ examinations impacted the teaching/learning process?
2. How has using only teacher-developed MCQ examinations impacted the teacher/student relationship?
3. How are students assessed formatively in the theory component of nursing courses?

Part I Introductory Questions

1. What are your experiences with writing multiple-choice questions?
2. What are the positive aspects of using faculty developed multiple-choice question examinations?
3. What are the negative aspects?
4. Think back to the last test that you designed, were there any problems with the MCQ used?

Part II Exploring the impact on the teaching/learning process

1. How has using only one method to assess students in the theory component of nursing courses impacted the way you teach?
2. How has using only one method to assess students in the theory component of nursing courses impacted how students study?
3. In what ways are the current tests a reflection of the student learning outcomes?

Part III

Exploring the impact on the teacher/student relationship

1. What are your thoughts and/or feelings as you write MCQ examinations?
2. What are your thoughts and/or feelings during the actual test?
3. What are your thoughts and/or feelings during "test reviews"?
4. What is the purpose of test reviews?
5. How has using MCQ examinations impacted students with limited English proficiency or who were educated in foreign schools?
6. As you know, students have written lengthy comments on the course end evaluations about the MCQ examinations. What are your thoughts and/or feelings as you read these comments?

Part IV

Exploring methods of formative assessment

1. What methods of formative assessment are used in the theory component of the nursing courses?
2. How familiar are you with other methods of student assessment to evaluate learning or assign grades?

Conclusion

What additional information may help me to gain a full understanding of this phenomenon?

Closing Statement

I want to thank all of you for taking time from your busy schedules to speak to me today. I want to remind you that I will share a transcript of this focus group with you and you

will have the opportunity to make any corrections or comments before final data analysis. Do you have any additional questions for me at this time?

Appendix H: Examples from Codebook

My codebook was maintained electronically and on sheets of typewritten papers. The example here was typed into this appendix as I had originally written along the margins or used both the highlighter functions and an actual highlighter pen. This is the ongoing coding of the final theme of test reviews and its evolution into need for metacognition on the part of both students and faculty.