The Effectiveness of Multiple-Choice Test Items in Assessing the Clinical Judgment Abilities of Prelicensure Registered Nursing Students

by

Tracy D. Holt

Capstone project submitted to

American Sentinel University

August 19, 2019
AMERICAN SENTINEL UNIVERSITY

The Capstone Committee for the Capstone Project of

Tracy D. Holt

The Effectiveness of Multiple-Choice Test Items in Assessing the Clinical Judgment Abilities of

Prelicensure Registered Nursing Students

DNP Project Chair: Dr. Sandra Cleveland

Committee Member: Dr. Sarah Gabua

Date of Final Defense: August 19, 2019
Abstract

The purpose of this study was to evaluate the effectiveness of multiple-choice questions in assessing the clinical judgment abilities of prelicensure registered nursing students. Nurse educators are responsible for ensuring fairness in testing and evaluating the development of clinical judgment abilities of students throughout the nursing program of study. This study compared multiple-choice test items developed using the current NCSBN® NCLEX-RN test item writing guidelines to multiple-choice test items developed using the new guidelines proposed by the NCSBN® Next Gen NCLEX-RN® project. The results indicated a statistically significant difference between the two types of multiple-choice items in the measurement of clinical judgment. Participants answered all the items developed according to the current NCSBN® NCLEX-RN test plan correctly within each of the five physiological categories 41% of the time, as compared to only 14% of the time on the items written according to the recommendations of the NCSBN® Nursing Clinical Judgment Model in the Next Gen Project. These findings indicate the participants understood the theoretical knowledge of the physiological concepts but could not apply the theoretical knowledge to the clinical scenario presented in the multiple-choice questions created using the NCSBN® Nursing Clinical Judgment framework. The information gained from this study provides evidence of a process that nurse educators can follow using the NCSBN® Nursing Clinical Judgment model as a guide to create multiple-choice test items to assess the clinical judgment level of nursing students.

Keywords: multiple-choice questions, clinical judgment, NCSBN® Next Gen project
Acknowledgments

I would like to express my sincere appreciation to God for allowing me the privilege of participating in the DNP program. Without His love, provision, and guidance the completion of this Capstone Project would not have been possible.

To my husband Scott who has supported me every step of the way. Thank you for believing in me. Your encouragement and faith in my abilities mean more to me than you will ever know.

To the practitioners and educators who served as subject matter review experts providing feedback in the item development process, thank you for taking time out of your busy schedules. Your expertise improved the quality of the items used in this project and is greatly appreciated.

Thank you to the Dean and faculty of the college where this study was conducted for allowing me access to your students. Thank you to the students who volunteered their time. I could not have conducted the study without your assistance and participation.

A sincere thank you to Dr. Kris Skalsky for her guidance in the data analysis of the findings of this project. Without your assistance, I would not have been able to complete the analysis and feel confident in the statistical findings of this study.

Thank you to my project committee member, Dr. Sarah Gabua, and my project chair, Dr. Sandra Cleveland. Your guidance, knowledge, and expertise inspired me to seek excellence throughout the development and completion of this project. However, I am most thankful for your kindness and compassion during the times when I wanted to give up. Your encouragement kept me going.
# Table of Contents

SECTION 1: INTRODUCTION .............................................................................................................1
  Background of the Problem........................................................................................................4
  Review and Summary of Relevant Literature ..........................................................................15
  Statement of the Problem ........................................................................................................30
  Purpose of the Project ...............................................................................................................32
  Significance of the Project ........................................................................................................34
  Nature, Scope and Limitations of the Project ..........................................................................35
  Theoretical Framework .............................................................................................................46
  Definition of Terms ..................................................................................................................55
  Summary ..................................................................................................................................58

SECTION II: METHODS ..................................................................................................................60
  Introduction ...............................................................................................................................60
  Project Design ..........................................................................................................................60
  Sample and Setting ..................................................................................................................62
  Instrumentation ........................................................................................................................68
  Data Collection ........................................................................................................................73
  Data Analysis Methods ............................................................................................................75
  Data Management Methods .....................................................................................................82
  Ethical Considerations .............................................................................................................83
  Internal and External Validity ....................................................................................................84
  Summary ..................................................................................................................................85

SECTION III: RESULTS AND DISCUSSION OF FINDINGS .....................................................87
  Introduction ...............................................................................................................................87
  Summary of Methods and Procedures .....................................................................................88
  Summary of Sample and Setting Characteristics ..................................................................95
  Major Findings ........................................................................................................................98
  Implications for Nursing Practice ..........................................................................................103
  Recommendations ...............................................................................................................104
  Recommendations for Future Research ................................................................................105
  Discussion ...............................................................................................................................106

Electronic copy available at: https://ssrn.com/abstract=3634804
Conclusion and Contributions to the Profession of Nursing ..................................................... 108
References ....................................................................................................................................... 110
Table 1 ........................................................................................................................................ 126
Table 2 ........................................................................................................................................ 127
Figure 2 ....................................................................................................................................... 129
Figure 3 ....................................................................................................................................... 130
Figure 4 ....................................................................................................................................... 131
Figure 5 ....................................................................................................................................... 132
Figure 6 ....................................................................................................................................... 133
Appendix A .................................................................................................................................... 134
Appendix B .................................................................................................................................... 135
Appendix C .................................................................................................................................... 136
Appendix D .................................................................................................................................... 137
Appendix E .................................................................................................................................... 148
Appendix F .................................................................................................................................... 149
Appendix G .................................................................................................................................... 155
Appendix H .................................................................................................................................... 157
Appendix I .................................................................................................................................... 160
Appendix J .................................................................................................................................... 161
Appendix K .................................................................................................................................... 162
Appendix L .................................................................................................................................... 163
Appendix M .................................................................................................................................... 164
Appendix N .................................................................................................................................... 165
Appendix O .................................................................................................................................... 166
Appendix P .................................................................................................................................... 167
Appendix Q .................................................................................................................................... 168
Appendix R .................................................................................................................................... 169
Appendix S .................................................................................................................................... 170

Electronic copy available at: https://ssrn.com/abstract=3634804
List of Tables

Table 1. Demographic Data.................................................................................................................. 125
Table 2. SPSS Output Chi Square for Independence Test................................................................. 126
List of Figures

Figure 1. Demographic Data................................................................. 89
Figure 2. Total Number of Competencies per Test Item Category.................. 128
Figure 3. Percentage of Items Correct per Nursing Process Category............... 129
Figure 4. Comparison of Percentage of Items Correct per Item Type............... 130
Figure 5. Item Level of Difficulty........................................................... 131
Figure 6. Point Biserial Discrimination Index of Test Items.......................... 132
Figure 7. Gender of Sample Participants.................................................. 95
Figure 8. Age Groups of Sample Participants.......................................... 95
Figure 9. Racial Groups of Sample Participants......................................... 96
Figure 10. Cumulative GPA of Sample Participants................................... 96
Figure 11. Previous Healthcare Education of Sample Participants................ 97
Figure 12. Participants Achieving Clinical Judgment Competency per Physiological Category................................................................. 99
SECTION 1: INTRODUCTION

Registered nurses are the members of the interdisciplinary healthcare team responsible for monitoring patients’ conditions throughout the hospital stay. Early recognition of problematic cues and effective implementation of evidence-based nursing interventions have been identified as key factors in preventing patient injury and death (Jessee, 2018). Of great concern is the increasing number of patients who suffer irreversible harm or death from preventable medical mistakes in hospitals throughout the United States in recent years (Makary & Daniel, 2016). Registered nurses (RNs) must possess the ability to accurately interpret assessment data and formulate sound clinical judgments to maintain patient safety to promote positive health outcomes (Kavanagh & Szweda, 2017).

The National Council of State Boards of Nursing (NCSBN®) is the regulatory body responsible for ensuring patient safety and public protection (NCSBN, 2010). Newly licensed registered nurses are held to the same scope of practice as experienced registered nurses and must provide safe, competent care to prevent patient injury (Purling & King, 2012). New RN graduates must be competent in recognizing early cues of clinical decline, analyzing the patient’s assessment findings, and formulating the correct clinical judgments to implement evidence-based interventions to prevent harm.

To monitor the types of challenges novice registered nurses encounter during the first year of employment, the NCSBN® conducts a practice analysis study every three years (NCSBN, 2015a). Information obtained from these studies is used as the basis for the development of the licensure test plans to ensure the licensing examinations remain up-to-date with current nursing practice (NCSBN, 2015b). The results of the 2014 NCSBN® Practice
Analysis Report indicated the majority of tasks performed by entry-level registered nurses require the ability to utilize sound clinical judgment (NCSBN, 2018, January).

The NCSBN® is revising the Nursing Council Licensure Examination for Registered Nurses (NCLEX-RN®) examination to ensure accurate evaluation of the clinical judgment abilities of novice registered nurses necessary to provide safe patient care. The NCSBN® has implemented the Next Generation Project to improve the test items on the NCLEX-RN® examination with the goal of accurately assessing the candidates’ ability to formulate safe clinical judgments (NCSBN, 2017, Fall). Dickison et al. (2016) have been tasked with developing a framework to use to create NCLEX-RN® test items capable of accurately assessing clinical judgment (p. 1). Building on the findings of Muntean (2012), Dickison et al. (2016) developed the NCBSN® Nursing Clinical Judgment (NCJ) Assessment Model, a multilayer assessment model capable of evaluating higher-order cognitive constructs using clinical case study information (Appendix A).

Evidence-based assessment methods are needed to evaluate student understanding of higher-order cognitive constructs to prepare RN students to pass the NCLEX-RN® examination and provide safe patient care. Bristol, Nelson, Sherrill, and Wangerin (2018) found there is a lack of evidence in the current nursing literature on the best practices for test item development, administration, and analysis. Exploration of evidence-based strategies to construct valid and reliable multiple-choice test items is relevant to current practices in nursing education.

Nurse educators are responsible for creating learning activities that promote student engagement in active learning strategies requiring demonstration of the application of theoretical knowledge in clinical practice (National League for Nursing [NLN], 2013). This is difficult to accomplish due to the size of registered nursing class cohorts, the fast pace of the undergraduate
curriculum, and the difficulty of developing accurate assessment methods. Undergraduate registered nursing programs continue to struggle to maintain qualified faculty due to the ever-increasing nurse educator shortage. Many programs have filled these spaces with practicing RNs who bring a wealth of clinical knowledge to the program but lack the training necessary to develop high-quality learning assessment methods (Su, Osisek, Montgomery, & Pellar, 2009; Tarrant, Knierim, Hayes, & Ware, 2006). Nursing faculty need an evidence-based process to follow to create assessment tools that are feasible to implement and analyze.

The purpose of this study was to test the validity of a framework that can be used to create multiple-choice test items that accurately assess the clinical judgment abilities of prelicensure registered nursing students. The analysis of statistical findings from individual student tests scores were beneficial in identifying students who are performing below the application and analysis cognitive operational levels required to formulate clinical judgments. The information can be used to assist students to identify learning activities that will facilitate development of higher-order cognitive thinking.

The researcher has organized this paper into three sections. Section one provides an introduction of the overview of this capstone project and the relevance of this topic in nursing education. A discussion of the background of the scope and prevalence of the identified problem provides a concise statement describing the problem and the significance of this study at this time. A summary of the findings of the review of the literature is provided to clearly define the purpose of this project, delineating the nature, scope, and limitations. Orlando’s Nursing Process Theory was used as the theoretical framework (Schmieding, 2002). A definition of terms is provided at the conclusion of Section one for clarification of terms.
Section two focuses on the methods used to complete this study. A description of the project design, the inclusion, and exclusion of the sample, and the setting in which the study took place are provided. The instruments used in the data collection and analysis are explained and a copy of each provided in the appendices of this paper. Ethical considerations during the implementation of this study, as well as the data management methods of the results of this study, are explained. Explicit documentation of plans for each step of this project and the how the results were analyzed are discussed, including the steps undertaken to facilitate internal and external validity.

Section three focuses on the results of the study and includes a discussion of the findings. A detailed description of the methods and procedures used to conduct the study are provided to promote replication by future researchers. Descriptive characteristics of the demographical information of the sample and the setting in which the study was conducted are included. An analysis of the major findings and the implications of these findings to nursing education and the contributions of these findings to the profession of nursing are conferred.

**Background of the Problem**

The NCSBN® was established 40-years ago and is responsible for maintaining public safety by ensuring graduates of nursing programs are competent to provide care to the citizens of the United States. Candidates are required to successfully complete the National Council Licensure Examination (NCLEX®) examination to obtain a license to practice (NCSBN, 2018b). The emphasis of the NCLEX for Registered Nurses (NCLEX-RN®) examination is to measure minimal competencies necessary to function as an entry-level novice RN in various healthcare settings including, but not limited to, acute medical-surgical-telemetry hospitals, community clinics, emergency departments, and home health/hospice agencies (NCSBN, 2015b). Over the
past decade the level of acuity of clients in the acute care hospitals has increased (Welton, 2017). The NCSBN® licensing examination requires frequent updates to accurately measure the competency of RN candidates.

Due to the enhancements in healthcare and technology, ongoing research and revisions are necessary to ensure the NCLEX-RN® examination remains relevant in the current healthcare environment. In 2012, the NCSBN® Examination Committee members questioned whether the NCLEX-RN® examination was accurately measuring the candidates’ clinical judgment ability (Grossenbacher & Kappel, 2018, Winter). This question sparked an inquiry into current nursing practice by the NCSBN®.

The investigation process began by commissioning a review of the literature to determine the current state of practice in nursing and testing in nursing education programs (Muntean, 2012). The findings from Muntean’s review of 200 peer-reviewed manuscripts were “alarming” and indicated that “50 percent of nurses were involved in errors in some way, and 65 percent of those errors were attributed to poor clinical decision-making skills” (Grossenbacher & Kappel, 2018, Winter, p. 11). These findings are of significant concern in correlation to research indicating medical errors are a preeminent cause of injury and death in the United States (Institute of Medicine [IOM], 1999; IOM, 2001; Makary & Daniel, 2016).

A somewhat encouraging finding from the review of the literature by Muntean (2012), is nurse educators have identified critical thinking, clinical decision making and clinical judgment as essential components to include throughout the nursing curricula (Grossenbacher & Kappel, 2018, Winter). Due to these findings, the NCSBN® Examination Committee members are exploring how to accurately assess the competency level of NCLEX-RN® examination
candidates in making safe clinical judgments as this is a public safety issue (Grossenbacher & Kappel, 2018, Winter).

**NCSBN® Strategic Practice Analysis Process**

The NCSBN® Strategic Practice Analysis study conducted from 2012-2014 evaluated the responsibilities of the entry-level RN (NCSBN, 2018, January). This extensive process consisted of environmental scans and site visits with direct observation of novice entry-level RNs performing patient care. Additionally, a job analysis survey of more than 2,000 entry-level RNs, four focus groups, and six workshops with more than 65 RN job experts were also conducted. The goal of this study was to determine the abilities required by entry-level RNs to provide safe, competent patient care to guide the development of the NCLEX-RN® examination. Of note, the observational portion of this practice analysis process was an unprecedented new strategy initiated by the NCSBN® in an attempt to accurately understand the true nature of the current nursing practice environment and the challenges thereof (Grossenbacher & Kappel, 2018, Winter; NCSBN, 2018, January).

The findings of the NCSBN® Strategic Practice Analysis Report indicated that attributes of problem solving, critical thinking, and clinical judgment are linked to the majority of the nursing tasks and skills required for an entry-level RN to provide safe patient care (NCSBN, 2017, Fall). Recommendations cited in the summary of the NCSBN® Strategic Practice Analysis Report (2018) included updating the existing NCLEX-RN® examination by developing new item types that accurately measure the candidate’s clinical judgment abilities.

**National Council Licensure Examination for Registered Nurses**

The NCSBN® is responsible for creating the NCLEX-RN® examination (NCSBN, 2018a). Upon review of the findings of the NCSBN® Strategic Practice Analysis Report
conducted from 2012 – 2014, the NCSBN® NCLEX Examination Committee members recommended an analysis of the current test items on the NCLEX-RN® examination to evaluate if the test items accurately evaluate clinical judgment (Grossenbacher & Kappel, 2018, Winter). The results revealed “large gaps” in the current test items that did not measure the candidate’s clinical judgment, critical thinking and problem-solving abilities (Grossenbacher & Kappel, 2018, Winter, p. 12). Consequently, the NCLEX® Examination Committee members determined substantial revisions of the test items were necessary for the NCLEX-RN® examination to be a valid means of assessment to ensure public safety.

The NCLEX® Examination Committee members commissioned research to determine the best evidence-based processes to ensure the licensing examination accurately measures safe clinical judgment abilities of the NCLEX-RN® examination candidates. This research project is known as the Next Generation Project (NCSBN, 2017, Fall). After extensive exploration, “a multilayer assessment model” was created to “establish a psychometric representation of nursing clinical judgment” called the NCSBN® NCJ Assessment Model (Dickison et al., 2016, p. 4). The multilayer design of this assessment model facilitates the design of test items that measure “cognitive subcomponents of NCJ” and to “identify at which process nurses may make errors” (Dickison et al., 2016, p. 3). Prototype test items were constructed and added as a research section to the NCLEX-RN® examination from October 2017 to April 2018 to test the reliability and validity of the items in measuring clinical judgment (Grossenbacher & Kappel, 2018, Winter). Preliminary test item analysis results indicate the items constructed following the NCSBN® NCJ Assessment Model accurately assess clinical judgment (P. Dickison, personal communication, March 23, 2018).
While this is exciting news, the design of these new test items is very different from the previous types of test items described in the item writing section of the detailed NCLEX-RN® examination test plan for nurse educators (NCSBN, 2015b). Evidence-based test item development tools are needed to assist in the development of valid and reliable test items to assess student understanding of critical concepts necessary to provide safe patient care. Historically there is a decline in pass rates when revisions to the NCLEX-RN® examination are implemented (NCSBN, 2003). Nurse educators have a responsibility to prepare student learning activities and assess student learning outcomes to adequately prepare undergraduate registered nursing students for the licensing examination.

**Impact on Nursing Education**

Nurse educators are responsible for designing learning strategies that facilitate the development of nursing students into competent registered nurses who can pass the NCLEX-RN® examination and provide safe patient care (NLN, 2013). To achieve this level of competency, nursing students must be able to process information and formulate sound clinical judgments. Nurse educators are tasked with the responsibility of developing assessment strategies that accurately evaluate each student without the provision of evidence-based, proven assessment tools.

While there is evidence in the current literature of clinical judgment assessment tools to assess student performance in the simulation and clinical environments (Manetti, 2017), there is a lack of evidence on proven strategies of how to develop test items that can be administered in the didactic nursing educational setting that accurately measure clinical judgment (Bailey, Mossey, Moroso, Cloutier, & Love, 2012; Bristol et al., 2018). The assessment methods used in the clinical and simulation educational settings rely on written assignments followed by a
discussion with the instructor using the Socratic Method to probe the student’s understanding of
the information and guide in the development of clinical judgment skills. This is not possible in
the classroom due to the limited amount of time and the size of the cohort.

Accurate assessment of large numbers of nursing students in undergraduate nursing
programs with methods that are easy to implement and provide useful analytical data on each
student is essential. Properly constructed examinations using either optical scanning analytic
software or computerized examination software are feasible. However, evidence of the
reliability of proven tools to develop valid exams items capable of examining higher cognitive
levels to determine if students grasp the complexity of the cognitive functions necessary for safe
practice is needed (Tarrant et al., 2006).

The NCSBN® NCJ Task Model provides a framework for the development of multiple-
choice test items to assess higher-order cognitive domains (Dickison et al., 2016). The
NCSBN® NCJ Task Model was used as the framework in developing the test items for this
study as this is the framework proposed for the development of the new NCLEX-RN®
examination test items. By developing items similar to actual NCLEX-RN® test items students
were exposed to the proposed changes on the NCLEX-RN® examination which is a
responsibility of undergraduate nursing education programs (Bristol & Brett, 2015). Most of the
test items on the NCLEX-RN® examination are multiple-choice test items, and multiple-choice
test items are a prevalent method of student assessment in undergraduate nursing programs
(Sutherland, Schwartz, & Dickison, 2012; Tarrant & Ware, 2008). This study provides evidence
of the process future nurse educators can use to construct high-quality multiple-choice test items
to utilize as an assessment method in undergraduate nursing programs.
Lack of nurse educator training. Due to the nurse educator shortage, nursing programs often seek clinical Masters prepared nurses to fill faculty vacancies (Su et al., 2009). While these nurses bring a wealth of clinical knowledge, they often have had little or no training regarding teaching and assessment. Even nurse educators who completed Masters level programs of study in nursing education often have difficulty developing high-quality multiple-choice test items (Tarrant et al., 2006). Nursing faculty frequently rely on items from textbook test banks. These test bank items have been proven to contain numerous errors and lack validity in testing higher-cognitive levels (Bristol et al., 2018; Masters et al., 2001; Tarrant et al., 2006). Information that is available regarding test item construction lacks empirical evidence verifying the reliability of these guidelines (Sutherland et al., 2012).

NCLEX-RN® examination changes. Nurse educators are responsible for assessing the level of understanding of the theoretical knowledge of each student, as well as the student’s ability to apply the theoretical knowledge in the patient care clinical setting (NLN, 2013). Verifying student comprehension of the content presented in the courses throughout the curricula in an undergraduate registered nursing program has always been a vital component in nursing education. Nursing education was originally conducted by apprenticeship in the patient care setting and the assessment of competency was conducted by direct observation. Nightingale established the first school of nursing in 1860. The curriculum consisted of both theoretical experiences in which students were required to attend lecture and take tests, and practical experiences in hospitals where competency continued to be assessed by direct observation (Ellis & Hartley, 2001).

In the mid-1900s nursing education programs began to transition from hospital-based programs to institutions of higher education. In the academic setting, the theoretical portion of
instruction was greatly increased (Ellis & Hartley, 2001; Zerwekh, 2018). Over the years the types of assessment have changed. Assessment by examination became, and remains, the primary mode of evaluation of student competency. While a clinical component has been maintained, the principal determining factor impacting student progression is theoretical examinations (Killingsworth, Kimble, & Sudia, 2015; Oermann, Saewert, Charaskika, & Yarbrough, 2009).

The first national examination required to obtain a nursing license was administered in 1941 (Johnson & Kappel, 2014, Spring). Over the years the NCLEX-RN® examination has changed dramatically. Prior to the invention of the computer, testing was conducted using paper and pencil which was time consuming to construct, administer, and grade. Today, the NCLEX-RN® examination is administered via computer.

In addition to the changes in the delivery format of the NCLEX-RN® examination, the structure has also changed. Initially, test construction was based on the different fields of nursing such as adult medical/surgical, pediatrics, obstetrics and gynecological, and psychiatric nursing. The candidate took an exam on each section and individual scores were generated. If a candidate was unsuccessful in one area, then the candidate was required to retake that section. In the 1980s test construction changed to a more concept-based structure. One score was generated for the entire exam indicating success or failure. If a candidate was unsuccessful, the candidate was provided with a breakdown of percentages for each concept, but the candidate was required to retake the entire exam.

The latest change to the structure of the exam differentiated the constructs into client’s needs categories. Today, the NCLEX-RN® examination is a complex, computerized adaptive test that uses extensive algorithms to evaluate the clinical judgment ability of the candidate. The
addition of the adaptive functionality of computerized testing has taken computerized testing to a
new level, allowing the test to adapt to the individual test taker thereby performing a more
thorough assessment of the student’s understanding (Johnson & Kappel, 2014, Spring; NCSBN,
2015b).

With the upcoming changes in test item format proposed by the Next Generation Project
(NCSBN, 2018a), the information regarding how to write test items that is currently available for
nurse educators does not align with the new NCLEX-RN® testing framework. Current test item
writing guidelines recommend developing a concise stem, containing only information about one
concept, devoid of any extraneous information (Billings & Halstead, 2016; NCSBN, 2015b;
Oermann & Gaberson, 2017). This is very different from the new NCLEX-RN® examination
item development process.

The NCSBN® NCJ Task Model provides a framework for developing short vignettes
containing patient data regarding two or more concepts the student must analyze to form a
hypothesis and generate the appropriate solution (Appendix B). All information, including the
answer choices, must realistically present patient situations and nursing interventions to evaluate
if the student can make a safe clinical judgment. In addition, multiple “sibling” items must be
developed to tests the same concept to validate the student can consistently formulate sound
clinical judgments thus eliminating threats to internal consistency from guessing (Dickison et al.,
2016, p. 12). The stem must contain information regarding at least two concepts for the student
to demonstrate the ability of cue recognition which is the foundation of the decision-making
process (Muntean, 2012). These types of items are more difficult to write than short-stemmed
items. This is concerning due to the level of poor quality of current practices in test item
development and the lack of evidence-based standards in testing and evaluation methods in nursing education today (Bristol et al., 2018; Tarrant et al., 2006).

**Responsibility to students.** Nurse educators have an ethical and legal responsibility to develop high-quality, valid examinations that ensure fairness and equity in testing (Killingsworth et al., 2015; Tarrant et al., 2006). Examinations are used throughout undergraduate nursing programs and determine continued student progression through the program (Tarrant et al., 2006). By examining the use of the NCSBN® NCJ Task Model as a framework for test item construction, this study hopes to provide future nurse educators with empirical evidence of the validity of this tool. The development of high-quality test items that accurately assess the clinical judgment abilities of nursing students in prelicensure nursing programs will be useful for future nurse educators in the construction of examinations that facilitate the assessment of student learning outcomes in didactic courses.

**Needs of the research organization.** This study was conducted at a community college in the Pacific Northwest. The community college offers an Associates in Nursing Direct Transfer Agreement/Major Related Program (DTA/MRP) degree. Students complete 135 credits at the community college and are then eligible to take the NCLEX-RN® examination and to transfer to a university to complete a bachelors in nursing (BSN). The community college implemented this change in Spring 2019 from the previously offered two-year Associate of Applied Science (AAS) degree to promote the Institute of Medicine’s goal of increasing the number of BSN prepared nurses to 80% by the year 2020, which is also a goal of the Washington State Nurses Association’s Academic Progression in Nursing (APIN) project (IOM, 2010; Sikma, 2015, November 6).
Faculty are concerned about the upcoming changes to the NCLEX-RN® announced by the NCSBN® in the Next Gen project (NCSBN, Fall, 2017). Over the past five years, the nursing faculty at this community college have implemented changes resulting in a 23 percent increase in the NCLEX-RN® pass rate. While the NCSBN® has not announced an implementation date for the upcoming changes, faculty feel changes in teaching and testing strategies should be incorporated into the current curriculum now to prepare the present cohort since these students will take the NCLEX-RN® examination in approximately two years. However, the faculty are unsure exactly what types of changes are needed to prepare students to answer the new type of items described in the Next Gen Project (NCSBN, Fall, 2017). The nursing faculty will use the results of this study to guide recommendations for changes in test item development and test analysis policies.

Writing valid and reliable test items is difficult and time consuming (Bristol et al., 2018). The research organization where this study was conducted does not require students to purchase access to a standardized testing program that provides a test bank from which faculty can develop examinations. The majority of the faculty use the test banks that are provided with the textbooks but tweak these items due to the accessibility of these items by the students on the world-wide-web (Madara et al., 2017). However, none of the test banks have been updated with the new NCJ format. Faculty are seeking information on how to write test items that are similar in structure to the new item type described in the Next Gen Project using the NCJ Task Model to adequately prepare students to pass the NCLEX-RN® examination (Dickison et al., 2016).

The purpose of this study was to evaluate the effectiveness of multiple-choice test items developed using the NCSBN® NCJ Task Model in accurately assessing the clinical judgment abilities of prelicensure registered nursing students as compared to the results from multiple-
choice test items developed without the use of the framework. The findings from this research will add to the current nursing education body of knowledge on test item construction and the measurement of clinical judgment in the didactic nursing educational setting.

**Review and Summary of Relevant Literature**

The purpose of the literature review was to examine current findings of peer-reviewed publications for any studies that have been conducted regarding multiple-choice test items and assessing the clinical judgment abilities of prelicensure registered nursing students. The searched was conducted to determine if any research has been done that:

1. Evaluated how well multiple-choice test items assess the clinical judgment abilities of prelicensure registered nursing students?
2. Established best practice guidelines for multiple-choice test item writing for nurse educators in prelicensure registered nursing programs?
3. Defined current practices of test item development in prelicensure registered nursing programs?

This section identifies the search strategies used for the literature review and a summary of the findings. Significance of the findings to nursing education and the limitations of the results were examined using Orlando’s Nursing Process Theory (Schmeiding, 2002).

**Search Strategy**

ProQuest, Cumulated Index for Nursing and Allied Health Literature (CINAHL), Ovid, and PubMed databases were searched using various combinations of the terms “prelicensure registered nursing students,” “undergraduate registered nursing programs,” “prelicensure registered nursing programs,” “clinical judgment,” “clinical reasoning,” “critical thinking” “multiple-choice test items,” “multiple-choice questions,” “test item development” and “testing
in nursing”. Search parameters were limited to peer-reviewed articles published in English conducted within the United States. Studies from non-nursing disciplines were excluded due to the unique requirements of prelicensure registered nursing education programs. A date limit was not applied.

The following narrative presents the relevant information obtained regarding test item development within the identified population, the intervention, and the outcome of this study, i.e., prelicensure registered nursing programs, multiple-choice test items, and clinical judgment, respectively. A synthesis of the findings is provided in each section addressing assessment practices in undergraduate RN educational programs, item writing and analysis guidelines, evidence-based test item development tools. A summary of the main themes and gaps identified with the recommended contributions to the field of study is also provided at the conclusion of this section.

Assessment in Prelicense Registered Nursing Education

Nurse educators are responsible for creating meaningful examinations that thoroughly evaluate the student’s level of understanding, and for administering the examinations in a fair, unbiased manner (Bristol et al., 2018; Clifton & Schriner, 2010; Killingsworth et al., 2015). Numerous assessment techniques are available to nurse educators but require adequate training to successfully implement. RNs practicing in the clinical setting are often recruited to fill vacant nurse educator positions due to the current nurse educator shortage (Tarrant et al., 2006). These clinical nurse educators usually have not had any formal training in learning assessment techniques.

Oermann et al. (2009) analyzed the results of a 29-item survey completed by 1,573 nursing faculty from prelicensure nursing education programs regarding assessment and grading
practices. Faculty reported using numerous forms of assessment strategies such as written papers (65%), group projects (63%), case study analysis (62%), presentations (49%) and reflective journals (40%). Examination methods consisted of the use of standardized tests (51%), teacher-made tests with varied types of items (61%), and teacher-made tests with multiple-choice test items only (55%). Fifty percent of the respondents reported the type of evaluation method used was determined by the amount of time needed to develop, implement and grade the assessment activity. A significant finding from the results is only fifty percent of the respondents considered research on the effectiveness of the assessment strategy selected to be important.

Killingsworth et al. (2015) conducted a descriptive correlational study of 127 BSN faculty in 31 states using a web-based survey to gather data regarding best practices for classroom testing. Faculty used a Likert scale of 1 (not skilled) to 7 (very skilled) to self-report their ability to construct tests (mean 5.3, $SD = 1.2$), perform item analysis (mean 5.5, $SD = 1.2$), and item revision (mean 5.6, $SD = 1.3$). The most common influence on test item construction was directly related to the course content (mean 6.7, $SD = .58$), with peer review of test items rated as least important (mean 4.2, $SD = 2.0$). The most frequently analyzed data to determine the validity of a test item was reported as the item difficulty level (mean 6.7, $SD = .76$). The overall results of the survey indicated faculty rated themselves as moderately to very skilled in the areas of item analysis and revision, with the lowest confidence self-reported in the area of test construction.

The impact of poorly constructed test items has far reaching implications on student success and patient safety. In their study of flawed test items, Tarrant and Ware (2008) found borderline students benefited from flawed exam items and would not have passed the exam if the flawed items would have been removed (p. 202). Masters et al. (2001) analyzed a random
sample of 2913 questions from 17 test banks that accompanied nursing text books and found 2233 errors. A study by Killingsworth (2015), revealed the majority of test items used in undergraduate nursing programs were selected from test banks that accompany the text books or were written by the nursing faculty.

Bristol et al. (2018) conducted a quantitative exploratory study using a survey consisting of 22 items about test item development, test administration, and test analysis practices of nursing faculty in prelicensure registered nursing programs. The sample consisted of 674 nursing faculty in undergraduate registered nursing programs in the United States. Results indicated a lack of consistency in test item development, administration, and analysis practices. Only 58% reported having a testing policy, but these policies primarily addressed academic dishonesty and how the examiners were to be proctored, not the development and analysis of the test items. Specific data regarding how test items were created indicated 69.4% of faculty modified items from test banks and 55.5% developed their own items. Only 8.8% reported strict guidelines regarding test item analysis (Bristol et al., 2018).

This study by Bristol et al. (2018) identified a gap in the implementation of evidence-based standards when developing and evaluating examinations. The researchers recommended further studies to identify evidence-based standards for test item construction, examination administration, and test item analysis practices in registered nursing educational programs. This is of great concern considering the findings of a number of studies indicating the frequency of errors in test bank items, and the lack of faculty who are adequately trained in test item development (Bristol et al., 2018; Masters et al., 2001; Su et al., 2009; Tarrant et al., 2006; Tarrant & Ware, 2012).
Multiple-Choice Test Items

The most common form of assessment in the didactic setting for large cohorts is examinations consisting of multiple-choice item types. Birkhead, Kelman, Zittel, and Jatulis (2018) conducted a descriptive correlational analysis study to evaluate the use of multiple-choice questions (MCQs) by nurse educators in RN programs in New York State. A 49-item survey tool was used to solicit data from 297 educators in 61 prelicensure RN programs. The results of the survey revealed MCQs constituted an average of 81% of the test questions on the examinations used throughout the prelicensure nursing programs. Fifty-four percent of the respondents indicated that 80% of the course grade was derived from the test scores. Chi-square analysis of the results indicated a statistically significant finding that in programs where a high percentage of MCQs are used on course examinations a high percentage of the overall course grade was calculated from the test scores ($\chi^2 = 65.03, p < .05$).

The advantages and disadvantages of using multiple-choice test items are well documented in the current literature. The advantages include: (a) the ease of use in a variety of settings, (b) the ability to assess a large population of students, and (c) the ease of grading and obtaining statistical analysis through the use of either computerized examination software or optical scanning software. However, to accurately evaluate the student’s clinical judgment abilities, the test items must be constructed to test higher-order cognitive domains. The major disadvantages are: (a) the difficulty in constructing the items, and (b) the amount of time required to develop the items (Masters et al., 2001; Tarrant et al., 2006). This is especially true of the type of multiple-choice test item necessary to measure the higher cognitive constructs of application, analysis, and evaluation.
Johnson (1999) designed a comparison study to evaluate how well multiple-choice test items measure clinical judgment. Using a 30-item exam containing questions similar to the items on the midwifery nursing certification examination, participants were asked to answer the questions and then rank the items on the degree to which each measured (a) midwifery management process, (b) stimulated analysis of the midwifery management process, (c) level of Bloom’s taxonomy, and (d) knowledge or clinical judgment. The sample of participants included 26 novice certified midwife nurses and 40 experienced certified midwife nurses. The researcher hypothesized that (a) the novice group would score higher on the items constructed at the cognitive level of knowledge as defined by Bloom’s, and (b) the experienced midwives would have higher scores on the items written at the cognitive levels of application and analysis required for clinical judgment.

Johnson (1999) used a one-tailed t-test to analyze the findings which indicated that the novice midwives did score an average of 1.08 points higher on the test items written at the knowledge level than the experienced group ($t[64] = 1.96, p < .05$). The findings, however, did not support the second hypothesis as the novice midwives scored an average of .39 points higher than the experienced midwives on the items written at the application and analysis level ($t[64] = 1.04, p > .05$). Johnson further evaluated the reliability of the judgment-type test items and found the reliability coefficient was low at .26, as compared to the knowledge-type test items reliability coefficient of .48. Johnson surmised that in light of these findings that the sample size of 66 participants may not be sufficient to accurately measure how well the items evaluate clinical judgment.

The format of the majority of the multiple-choice test items consisted of short-stemmed questions. One of the items provided patient data in an extended stem in the form of a short
vignette. Of note is that 70% of the participants ranked this long-stem vignette item as the highest in assessing clinical judgment. Johnson recommended the item format structure on the certification exam be changed to include realistic patient data in the stem of the item, followed by questions that require the test-taker to analyze the data and formulate an active decision to evaluate clinical judgment.

In order to evaluate the student’s understanding of critical concepts of patient care, information relevant to at least two concepts must be presented in the stem of the item. Research indicates the use of vignettes containing accurate clinical data requiring the student to analyze the data and formulate a clinical judgment decision is the best test item format to use (Case & Swanson, 2003; Coderre, Harasym, Mandin, & Fick, 2004; Dawson, Comer, Kossick, Neubrander, 2014; Dickison et al., 2016; Johnson, 1999; Tarrant & Ware, 2012). Several sibling items must be developed to test the same construct and examined using a linked statistical correlation to eliminate false positives from guessing (Dickison et al., 2016). Test items must present clinical decision-making tasks relevant to real-life clinical practice for the student to demonstrate the ability to formulate sound clinical judgments in clinical situations similar to what is required in practice (Su et al., 2009). This is very different from the current practice of creating a short-stemmed item with limited data (Billings & Halstead, 2016; NCLEX, 2016; Oermann & Gaberson, 2017; Sutherland et al., 2012).

Evidence from the literature review revealed a lack of faculty training on test writing techniques and frequent use of flawed test bank items (Bristol et al., 2018; Masters et al., 2001; Tarrant et al., 2006; White & Heitzler, 2018). Schroeder (2007) conducted a quasi-experimental comparison study to evaluate the impact of a test-taking workshop on registered nursing students’ critical thinking test scores. The sample consisted of 37 first semester students in an
associate degree nursing program. The convenience sample of students was divided into a control group \((n = 16)\) and an experimental group \((n = 21)\). Both groups completed a standardized pre-test. The experimental group was then exposed to a test-taking workshop. Both groups were then given the Critical Thinking Process Test.

Statistical analysis of the results of Schroeder’s study (2007) using a one-tailed \(t\)-test revealed the mean scores of the two groups were statistically insignificant using \(p < .05\). The \(t\)-test score of the means of all four multiple-choice examinations was .19, which was not considered statistically significant at the .05 level. The experimental group did not show improvement in their ability to apply critical thinking skills to examinations consisting of multiple-choice test items written at higher-cognitive levels. These results indicate that simply teaching students test-taking skills does not improve their ability to critically think and formulate safe clinical judgments. Schroeder recommended continued research in this area with larger samples sizes that are randomly selected to help define critical thinking as it applies to nursing and identify a reliable standardized test to measure critical thinking skills.

While text books identify standards for nursing examination development (Billings & Halstead, 2016; Oermann & Gaberson, 2017), findings of studies in the current literature indicate a lack of adherence to these standards in nursing education programs (Bristol et al., 2018; Clifton & Schriner, 2010; White and Heitzler, 2018). DePew (2001) conducted a descriptive exploratory study to evaluate how nursing faculty in associate degree nursing (ADN) programs and baccalaureate degree nursing (BSN) programs used multiple-choice test items. DePew designed a 62-item survey instrument to assess if faculty performed activities to determine item validity and reliability, as well as activities of overall test evaluation.
The sample consisted of 1,086 undergraduate nursing faculty from 329 nursing programs in the United States. Results indicated that 76% of ADN and BSN faculty performed activities to verify the validity of the items, with the ADN faculty having a slightly higher mean score of 52.46, than the BSN faculty 48.6. A total of 87% of the ADN and BSN faculty performed activities to determine the reliability of the items, with the BSN faculty having a slightly higher mean score of 44.67, than the ADN faculty 42.49. A reported 81% of ADN and BSN faculty conducted a test evaluation after each exam with the mean scores for each group almost equal at 93.91 for the AND faculty, and 93.44 for the BSN faculty.

DePew (2001) surmised a statistically significant correlation could not be determined between faculty validity, reliability, and test evaluation activities to the overall NCLEX-RN pass rates. However, a slight correlation was found in the standard deviation scores of item validity activities (3.82) and test evaluation activities (6.21) with the higher NCLEX-RN® pass rate groups indicating that consistency in these activities may have a positive impact on NCLEX-RN® pass rates. DePew recommends nursing faculty should attend formalized training on test item development, test administration and test item analysis, and resources should be made available to nursing faculty. DePew also acknowledges the lack of studies regarding test item development in nursing education and calls for repeat studies with larger samples due to the impact of multiple-choice examinations on the students’ progression through the nursing program.

Several peer-reviewed articles in the literature provide guidelines for test item construction, however, none of these provide empirical evidence validating the effectiveness of these guidelines (Bristol & Brett, 2015; Masters et al., 2001; Stanton, 1983; Sutherland et al., 2012; Tarrant et al., 2006; Tarrant & Ware, 2008). This is a significant concern given the impact
of the theoretical examinations on the student’s ability to progress through the program, successfully complete the NCSBN-RN® examination, and provide safe, competent patient care. The findings of this literature review revealed a gap in the literature regarding evidence-based test item construction tools (Bristol et al., 2018; Clifton & Schriner, 2010; DePew, 2001; Johnson, 1999; Killingsworth, 2013; Schroeder, 2007; White & Heitzler, 2018).

There is conflicting data in the current literature regarding the ability to measure higher-order cognitive constructs using multiple-choice test items. Masters et al. (2001) claimed multiple-choice test items can only be used to assess the lower levels of Bloom’s taxonomy. In contrast, Dickison et al. (2016), and Case and Swanson (2002) declared multiple-choice test items can accurately assess the higher cognitive constructs of Bloom’s taxonomy. In order to measure a construct, the contextual factors must be clearly defined. Numerous definitions of clinical judgment exist within the literature, however, none of the definitions provide clear defining criteria comprising the construct of clinical judgment that facilitate the measurement of the construct by written examinations (Cappelletti, Engel, & Prentice, 2014).

Several studies provide assessment methods of measuring clinical judgment in the clinical and simulation educational environments, but these methods consisted of one-to-one observation by the instructor, use of rubrics to evaluate the student’s actions, or written assignments followed by detailed written feedback and/or discussion with the instructor (Gillespie & Peterson, 2009; Jesse, 2018; Manetti, 2017). These types of assessment modalities are not feasible in the didactic classroom. The theoretical component of nursing education is typically conducted in traditional classroom settings by one instructor on one to two days a week for two to six hours. Instructors cannot conduct a one-to-one assessment of each student for every assignment and examination due to the time limitations.
Consequently, the NCSBN® has developed a model of the process by which nursing clinical judgment is cognitively constructed based on the information-processing theory (Dickison et al., 2016; Muntean, 2012). Using the definitions of each of these constructs, Dickison et al. (2016) developed the NCSBN® NCJ Task Model that provides a framework for test item development. The structure of the task model framework corresponds to the NCSBN® NCJ Assessment Model upon which analysis of the cognitive level of the answer to the test item can be evaluated. The structure of these new test items is drastically different from the current NCLEX-RN® test items (Dickison et al., 2016; NCSBN, 2015b).

The development of the new NCSBN® NCJ Task Model was the only tool identified from the literature review that provided an evidence-based framework for item development (Dickison et al., 2016). However, no studies have been conducted indicating if the NCSBN® NCJ Task Model can be utilized by nursing faculty to develop multiple-choice test items. The NCJ Task Model does support other studies indicating the best way to evaluate clinical judgment with multiple-choice test items is to provide the student with patient data in the item stem in the form of a vignette (Case & Swanson, 2003; Coderre et al., 2004; Dickison et al., 2016; Johnson, 1999; Tarrant & Ware, 2012).

Due to these findings and the upcoming changes to the NCLEX-RN® examination (Grossenbacher & Kappel, 2018; Winter NCSBN, 2017, Fall), the NCSBN® NCJ Task Model (Appendix B) was used as the framework for the development of the independent variables for the study, the multiple-choice test items. The NCSBN® NCJ Assessment Model (Appendix A) was used to analyze the results of the dependent variables of the study, the participants’ answer selections on the multiple-choice test items (Dickison et al., 2016; NCSBN, 2017, Fall). A table depicting the relationship of the NCJ cognitive operational levels to Orlando’s Nursing Process
and Bloom’s Cognitive Activities is provided in Appendix C (Anderson et al., 2001; Schmeiding, 2002).

**Clinical Judgment**

Nurse educators in prelicensure RN programs place great emphasis on the importance of assessing the clinical judgment abilities of RN students. Numerous studies addressed the impact of the RNs clinical judgment ability on providing competent patient care (Di Vito-Thomas, 2000; Jesse, 2018; Kavanagh & Szweda, 2017; NCSBN, 2015a; Purling & King, 2012). Del Bueno (2005) analyzed 10-years of data from a quantitative quasi-experimental study that assessed the clinical judgment abilities of experienced and novice registered nurses (RNs). The Performance Based Development System (PDBS) was used to evaluate 20,413 RNs with greater than 1-year experience, and 10,988 RNs with less than one-year experience in 350 health care agencies in 46 states. The initial assessment scores indicated only 35% of the RNs with less than 1-year experience met entry-level expectations of clinical judgment to perform safe patient care independently. After completing an individualized improvement plan, reassessment scores indicated 70% of these inexperienced RNs showed improvement in clinical judgment abilities. Del Bueno recommended nursing education programs should provide experience in visual simulations and increased clinical patient care hours for students to learn how to apply theoretical knowledge to patient care effectively.

Kavanagh and Szweda (2017) analyzed the results of 5,000 newly licensed RNs on the Performance-Based Development System assessment. The study was conducted over a five-year period and consisted of graduates from 140 nursing programs in 21 states. The findings indicated that only 23% of the newly licensed RNs were competent to perform patient care independently. The findings of this study were utilized by a medical center in the mid-west to
develop a residency program to assist newly licensed RNs to develop good clinical judgment prior to performing patient care independently. Kavanagh and Szweda (2017) recommended extensive curriculum changes in nursing education programs to assist students to develop the clinical reasoning skills necessary for safe practice.

While the ability to accurately analyze patient assessment data and formulate safe clinical judgments is well documented in the literature as a basic requirement of the registered nurse’s scope of practice, there is a lack of evidence-based studies on how to accurately assess these cognitive functions in the didactic learning environment (American Nurses Association [ANA], 2015a; Bristol & Brett, 2015; Bristol et al., 2018; Clifton & Schriner, 2010; Jesse, 2018; Killingsworth, 2013; Masters et al., 2001; Muntean, 2012; NCSBN, 2018, January; Stanton, 1983; Sutherland et al., 2012; Tarrant et al., 2006; Tarrant & Ware, 2008). Dawson et al., (2014) evaluated the effectiveness of the Script Concordance Test (SCT) in measuring the clinical reasoning skills of 48 first-year BSN students by comparing the students’ test scores to the scores of a panel of experts consisting of experienced RNs, a minimum of 10 experts for each area of clinical practice. The SCT consists of realistic clinical scenarios followed by different types of test items. The SCT for this study contained 29 clinical scenarios and 76 test items.

Cronbach alpha for the total test was 0.85 with a reliability factor of >0.8 indicating good internal consistency of the items. The students’ mean score on the overall exam (63.71%, \(SD = 8.5\)) was significantly lower than the expert panels’ mean score (78.52%, \(SD = 8.09\)). Cohen’s \(d\) of the effect size (\(d = 1.51\)) indicated a high significance between the two groups.

The findings of Dawson et al., (2014) validated the SCT as a reliable measure of overall clinical reasoning in nursing students. The limitations of the use of this tool are the amount of time and expertise required to develop realistic case scenarios and the recruitment of at least ten
experienced RNs for each area of practice that is tested who are willing to serve on the panel. This is not feasible for the multiple numbers of examinations required in each course throughout the undergraduate nursing curriculum.

Miller, Tedford, and Lehmann (1981) conducted a study to evaluate the clinical judgment abilities of registered nursing (RN) students. Students performed physical assessments of 469 clients between the ages of 45 and 93 at an outpatient clinic. The students were required to critically analyze the findings and determine if the client should be referred to a physician. The students completed a referral form with the abnormal findings which the client was to present to their personal physician for validation of the student’s findings, and any actions or treatments needed. These forms were then returned to the college of nursing via mail for statistical analysis. The results indicated the students’ clinical judgment was accurate in 78.8% of the cases referred to the physician for follow-up. The researchers hypothesized that the use of numerous personal physicians instead of one physician that was readily available for immediate follow-up had a negative impact on the findings. Miller et al. (1981) recommended a repeat study with a dedicated physician on-site to evaluate the abnormal findings, as well as to validate the normal findings, and offer immediate feedback to the RN students.

While the study conducted by Miller et al. (1981) provided an excellent example of how to facilitate the measurement of clinical judgment in the clinical educational environment, the literature review indicates evidence-based assessment modalities proven to accurately evaluate the clinical judgment abilities of registered nursing students in undergraduate theory classes are needed. Several studies provided strategies on how to assess clinical judgment abilities using simulation and also in the clinical educational setting (Johnson et al., 2012; Lasater, 2007; Manetti, 2017). However, there is a complete lack of studies on evidence-based proven
strategies on how to accurately assess the clinical judgment abilities of RN students in didactic prelicensure nursing courses using multiple-choice questions (Bristol et al., 2018; Killingsworth, 2013).

The study of the use of the SCT by Dawson et al., (2014) relies on the opinions of the members of the expert panel and does not provide specific criteria delineating the defining characteristics and measurement criteria to utilize to perform consistent measurement of clinical judgment with ever changing expert panels. The importance of assessing the clinical judgment abilities of prelicensure nursing students in undergraduate nursing programs is of great concern due to the following issues: (a) the student success rate on the NCLEX-RN® examination, (b) the impact of the first-time NCLEX-RN® examination pass rate percentages on the accreditation status of the nursing program, and, most importantly, (c) the ability of the candidate as a novice nurse to provide safe patient care (Bristol et al., 2018; Clifton & Schriner, 2010; Jesse, 2018; Muntean, 2012; NCSBN, 2018, January).

**Summary of the Review of the Literature**

This review of the relevant literature revealed a gap in the current nursing literature of proven evidence-based guidelines for test item writing and analysis in undergraduate nursing educational programs (Bailey et al., 2012; Bristol et al., 2018; Clifton & Schriner, 2010; Killingsworth et al., 2015; Masters et al., 2001; Tarrant et al., 2006). A significant weakness identified in the literature was the majority of studies that had been conducted focused on the measurement of clinical judgment in the simulation or clinical learning environments. The study by Killingsworth et al. (2015) revealed, the majority of “nursing education research has focused on clinical assessment and evaluation, with scant research focused on classroom assessment and evaluation” (p. 221).
One theme identified as a strength of the literature review was several qualitative studies consisting of feedback from surveys of nursing faculty that provided recommendations for writing test items (Bristol et al., 2018; Killingsworth, 2013; Killingsworth et al., 2015; Masters et al., 2001; Oerman, Saewert, Charaskika, & Yarbrough, 2009). However, no studies could be found validating the recommendations of test item development provided by nursing faculty by quantitative experimental research (Bristol et al., 2018; Haladyna, Downing, & Rodriguez, 2002; Masters et al., 2001; Morrison & Free, 2001; Oermann et al., 2009; Sutherland et al., 2012; Tarrant & Ware, 2012). White and Heitzler (2018) indicated that current examination methods focus more on the student’s test taking abilities than on accurately evaluating the student’s level of knowledge. The review of the literature indicated evidence-based item writing development tools are needed to guide construction of items that accurately assess the higher-cognitive constructs required for sound clinical judgments in prelicensure registered nursing educational programs.

**Statement of the Problem**

The literature review revealed a lack of evidence-based methods for writing multiple-choice test items capable of reliably assessing the higher-order cognitive processes required to formulate safe clinical judgments (Bailey et al., 2012; Clifton & Schriner, 2010; Masters et al., 2001). Registered nursing students in prelicensure RN programs depend on nurse educators to prepare them to pass the NCLEX-RN® examination and to competently function as novice RNs. Nursing faculty at the community college where this study was conducted desire information on how to update current course learning activities and assessment techniques to evaluate the development of the students’ ability to form competent clinical judgments.
Historically, evidence indicates changes in the NCLEX-RN® examination result in a decline in first-time pass rates (Alexander & Brophy, 1997; NCSBN, 2003; NCSBN, 2013). The NCSBN® is in the process of making major revisions to the NCLEX-RN® (Grossenbacher & Kappel, 2018, Winter; NCSBN, 2017, Fall). The nursing faculty at the research organization are concerned the current cohort will not be prepared to answer the new type of items measuring clinical judgment on the NCLEX-RN® resulting in a decline the programs pass rates. Nurse educators are responsible for identifying at-risk students and assisting these students to create an improvement plan to successfully complete the program and pass the NCLEX-RN® examination (Bristol et al., 2018; Tarrant et al., 2006). The results of this study will provide information regarding the usefulness of the NCSBN® NCJ Task Model as a framework, or lack thereof, to follow in the development of test items that assess the RN student’s clinical reasoning skills as well as how to identify the student’s level of understanding of NCJ.

Nurse educators are also responsible for protecting society at large by ensuring RN students demonstrate understanding of the critical concepts of clinical judgment before allowing the students to complete the program of study (NLN, 2013). Immediately upon achieving licensure, novice RNs are held to the same scope of practice as experienced RNs (Purling & King, 2012). Nursing faculty in undergraduate RN programs are responsible for assisting students to develop the knowledge and skills necessary to perform safe patient care as RNs. Novice RNs must recognize early signs and symptoms of patient decline, analyze the information, and implement the most appropriate intervention to promote optimal health and prevent harm (Jessee, 2018; Kavanagh & Szweda, 2017). By using the information provided from the research of Dickison et al., (2016) a process for test item construction and item analysis
that can be used by future nursing faculty in undergraduate programs to consistently develop test items that accurately assess the RN student’s clinical judgment abilities was explored.

The consequences of failing to adequately assess the clinical judgment abilities of RN students in undergraduate programs include (a) a negative impact on the student’s ability to pass the NCLEX-RN® examination and obtain gainful employment, (b) increasing the nursing shortage, and (c) poor patient outcomes from graduates who are unprepared to provide safe patient care as novice RNs. A possible solution proposed by this project was to establish an evidence-based method for developing multiple-choice test items to provide future nurse educators with an accurate method of assessment of the RN student’s cognitive operational level of clinical judgment. The results of this research add to the body of knowledge in the nursing education literature by providing evidence of the effectiveness of the NCSBN® NCJ Task Model as a framework to utilize to consistently develop multiple-choice test items that measure the cognitive level of clinical judgment of prelicensure registered nursing students.

**Purpose of the Project**

The purpose of this quantitative quasi-experimental comparative explanatory study was to examine the effectiveness of multiple-choice test items developed using the NCSBN® NCJ Task Model as a framework in measuring the cognitive level of understanding of prelicensure RN students. The objectives of this study were to determine if:

1. The NCSBN® NCJ Task Model is a framework that can be used to guide the development of multiple-choice test items (Dickison et al., 2016).

2. The student’s cognitive operational level of clinical judgment can be determined by examining the student’s answers to the multiple-choice test items created using the framework.
By performing this study this researcher sought to answer the following question: Is there a difference in the measurement of the nursing student’s cognitive level of clinical judgment when comparing the results of student’s answer selections to multiple-choice test items created using the NCSBN® NCJ Task Model multiple-choice test items created using the 2016 NCSBN® Test Plan? The PICOT formatted question was as follows:

- Population – Pre-licensure registered nursing students
- Intervention – Multiple-choice test items developed using the NCSBN® NCJ Task Model
- Comparison – Multiple-choice test items developed using the 2016 NCSBN® Test Plan
- Outcome – Accurate assessment of the cognitive function of nursing clinical judgment

The results were evaluated using the following hypothesis:

- Null Hypothesis: There are no statistically significant differences in the scores of the multiple-choice test items developed using the NCSBN® NCJ Model as a framework, as compared to the multiple-choice test items developed without using a framework.
- Hypothesis: There are statistically significant differences in the scores of the multiple-choice test items developed using the NCSBN® NCJ Model as a framework, as compared to the multiple-choice test items developed without using a framework.

The aim of this project was to establish evidence that multiple-choice test items created using the NCSBN® NCJ Task Model are effective in measuring registered nursing student’s level of understanding of a clinical judgment by evaluating the student’s ability to apply theoretical knowledge to an actual client scenario and make a safe clinical judgment regarding the nursing actions required to provide competent care for the client. Empirical evidence validating a process for evaluating NCJ in the didactic setting with large cohorts that does not
require lengthy, one-to-one discussions with each student as class time is not provided for this type of assessment is needed. Efficient assessment of large cohorts of RN students can occur by creating an examination using multiple-choice test items. Computerized testing software or optical scanning software can be used to facilitate quick grading and generate item analysis reports to determine the student’s level of clinical judgment for each clinical concept tested.

**Significance of the Project**

The project is significant to nursing education because this study examined the feasibility of the NCSBN® NCJ Task Model as a framework for test item development (Dickison et al., 2016). The findings add to the body of knowledge regarding measuring clinical judgment with multiple-choice test items which is lacking in the current literature. The findings are beneficial to future nurse educators in providing an evidence-based process for developing multiple-choice test items that accurately assess the RN student’s cognitive level that is necessary to formulate sound clinical judgments.

The future benefits of this study to students include (a) exposure to NCLEX-RN® examination-style questions, (b) identification of cognitive functional level, and (c) analysis of contextual factor patterns that influence the decision-making process in constructing clinical judgments. The findings from future individual student’s item analysis can be utilized by future nurse educators to advise the student on effective study strategies to improve in the identified areas of deficiency. The future benefit of the application of the findings of this study may improve the ability to provide individualized feedback to guide the student to focus on new learning activities that directly impact the problematic contextual factors identified from the item analysis, thereby improving the student’s ability to create safe clinical judgments.
A long-term benefit that is anticipated for the prelicensure nursing program where this study was conducted is a presentation of the item development process to teach faculty how to use the NCSBN® NCJ Task Model to develop multiple-choice test items. Incorporation of these items into future courses may be beneficial in preventing a decline in NCLEX-RN® pass rates with the upcoming changes to the NCLEX-RN® examination. First-time NCLEX-RN® pass rates of the graduates of the program is a key component in maintaining the program’s accreditation (Accreditation Commission for Education in Nursing, 2017; National League for Nursing Commission for Nursing Education Accreditation, 2016).

Successful completion of the NCLEX-RN® examination is required for the graduates of the program to obtain a license to practice as RNs (NCSBN, 2015b). Therefore, adequately preparing the RN students to pass the NCLEX-RN® examination and obtain gainful employment has a direct impact on relieving the RN shortage in the community the college serves (Clarke, 2016). Subsequently, the data provided in the test item analysis may aide nursing faculty in the identification of future at-risk RN students who lack the critical thinking skills necessary to formulate sound clinical judgments is a patient safety issue that affects both the local community and society at large by impacting the quality of care provided by the graduates of this prelicensure nursing program as novice RNs (Jessee, 2018).

**Nature, Scope and Limitations of the Project**

This section provides a discussion of the nature of this project, the instrumentation that was used, and how the data was collected and analyzed. The limitations and delimitations are also examined in this section. The scope of the project is expounded upon, as well as the feasibility of conducting the project in the proposed environment with the selected population.
Research Design

This prospective quantitative quasi-experimental study sought to determine the accuracy of multiple-choice test items developed using the NCSBN® NCJ Task Model as a framework in measuring the higher-order cognitive constructs of clinical judgment in prelicensure nursing students (Dickison et al., 2016). The NCSBN® (2018, Winter) defines the clinical judgment process as the “skill of recognizing cues about a clinical situation, generating and weighing hypotheses, taking action and evaluating outcomes for the purpose of arriving at a satisfactory clinical outcome” (p. 3). Two groups of test items were created to test the accuracy of multiple-choice test items in measuring clinical judgment.

The topics for the exam were the five physiological concepts of metabolism, perfusion, oxygenation, fluid and electrolytes, and infection. The researcher created three items for each of the five physiological concepts using the NCSBN® NCJ Task Model (Dickison et al., 2016). This group of items was labeled the framework group (Appendix D). The researcher also created three items for each of the five physiological concepts using the current NCSBN® 2016 Test Plan. This group of items was labeled the non-framework group. The exam consisted of a total of 30 items. Analysis of the cohort scores was conducted to determine if any statistically significant difference exists between the two groups of items.

The independent variable group, categorized the framework group, consisted of 15 test items developed using the NCSBN® NCJ Task Model which is based on the NCSBN® definition of clinical judgment. Each item stem contained patient assessment data for at least two physiological concepts. Each stem was followed by a series of three questions developed at the cognitive operational levels of Layer three and Layer two as indicated by the interval rating
layer scale on the NCSBN® NCJ Assessment Model (Dickison et al., 2016). This continuum measure of interval rating was used in the analysis of the student scores on each question.

The first item following the vignette required the participant to recognize and analyze the cues provided in the vignette to determine the priority nursing assessment that should be performed. This item was written at Layer three of the NCSBN® NCJ Assessment Model. Item two required the participant to prioritize hypotheses and generate solutions for the appropriate nursing action. This item was written at Layer two of the NCSBN® NCJ Assessment Model. Item three probed the participant's ability to evaluate outcomes. This item was written at Layer two of the NCSBN® NCJ Assessment Model.

The nominal scores of the participants answer selections of the categorical measures were analyzed to determine the student’s interval rating on clinical judgment ability continuum (Appendix E), per item, per concept and on the overall exam (Dickison et al., 2016). The participant must select the correct answer for all three of the framework items to demonstrate achievement of clinical judgment in the physiological concept. This is layer one of the NCSBN® NCJ Assessment Model.

The format for the construction of framework group items adhered to the recommendations of Dickison et al. (2016) and included the following:

1. The stem of the item consisted of realistic client assessment data pertaining to two (or more) clinical concepts in the form of a short vignette

2. The stem was followed by a series of questions that elicited the following active decision-making processes of cognition:
   a. Recognition of the cues presented in the assessment data
   b. Analysis of cues
c. Generation of hypotheses
d. Prioritization of hypotheses
e. Creation of correct nursing action to implement for top priority hypothesis
f. Evaluation of the outcomes of the nursing action implemented

3. A feasible answer choice for each concept presented in the stem
4. The correct answer choice represented a distinct level on the NCSBN® NCJ Assessment Model
5. The examination contained sibling items that test the student’s ability to accurately interpret cues of a physiological indicator of a critical change in client status (three items per concept) to prevent bias from guessing the correct answer on one item

The second group of test items, identified as the non-framework group (Appendix F), were created without the guidance of the NCSBN® NCJ Task Model. This group consisted of 15 test items developed using the guidelines from the NCSBN® 2016 NCLEX-RN® Detailed Test Plan for Nurse Educators (2016). Since the development of the test items for this project, the NCSBN® published the updated 2019 NCLEX-RN® Test Plan (2018b). Review of the updated 2019 NCLEX-RN® Test Plan did not reveal any new requirements for test item construction. The updates to the NCLEX-RN® Test Plan consisted of minor editorial changes, the addition, and reclassification of activity statements, and changes in NCLEX terminology (NCSBN, 2019), therefore, the items created using the 2016 NCLEX-RN® Test Plan remain relevant for use in this study.

The non-framework group test items covered the same five physiological concepts but were independent items without seriation of questions. The stems of these items only contained information regarding one concept. Each item was categorized using Bloom’s taxonomy of
cognitive domains (Anderson et al., 2001) as this is the classification method used in the NCLEX-RN® test plan (2016, p. 4). This categorical method of measurement was used to analyze the student’s scores on these items.

A blueprint of the complete exam (Appendix G) indicates the NCBSN® client’s needs category (NCSBN, 2015b), Bloom’s cognitive process (Anderson et al., 2001), Orlando’s Nursing Process (Doenges & Moorhouse, 2013), clinical concept, and any corresponding siblings for each test item. Subject matter experts (SME) reviewed all test items prior to use and ensured (a) the assessment data in the stem was accurate, (b) the actions delineated in the answer choices adhered to current practice standards, and (c) the rationales for both the correct and incorrect answer choices were theoretically sound.

Instrumentation

The examination was administered electronically using the Canvas® Learning Management System (LMS) in the nursing college computer lab. Canvas® is one of the world’s leading LMS companies with an established record of proven reliability in performance capabilities, accurate data collection, security of information, and maintaining the privacy of participants (Instructure, 2018). Canvas® is currently the LMS used in the nursing program. The students were familiar with the functions on the Canvas® quiz software.

A nursing instructor proctored the exam. Each participant was provided with a test student user identification (ID) and password to use to sign-in to the Canvas course that was created by the researcher during the Canvas course set-up to maintain confidentiality. These test student IDs did not contain any data that could be used to identify the participants. The students were provided with the examination access code by the proctor once all participants successfully signed in to the Canvas course. The IP addresses were filtered to prevent the proctor from being
able to identify the participant's results by the location of the computer used to access the Canvas course examination. The first screen contained the consent form (Appendix H) which the participants acknowledged prior to accessing the examination. A written copy of the consent form was also available for the participants. Each exam item was presented one at a time. An answer to each question was required prior to moving to the next item. Backtracking was not allowed. The final screen contained items for collection of demographical data. Access per participant was set at one-time only to automatically close immediately after the participant submitted the exam to maintain security.

Participants were encouraged to complete all test items. However, participants were allowed to close the exam at any time. The maximum time limit to complete the exam was one hour. Once the exam was submitted for grading, the correct answers and rationales appeared for independent review. After all participants completed the examination, the proctor conducted a review of the examination with all participants collectively to clarify any items to prevent confusion or misunderstandings.

Data Analysis

Statistical analysis reports were generated using the Canvas® reports function. The reports were void of any participant identification data to maintain anonymity. The NCSBN® NCJ Assessment Model (Dickison et al., 2016) was utilized to measure the outcome of clinical judgment (Appendix A). Descriptive and inferential statistics measures were used to complete a quantitative analysis of the data to determine if there was a difference in the measurement of the nursing student’s cognitive level of clinical judgment when evaluating the results of student’s answer selections to multiple-choice test items created using the NCSBN® NCJ Task Model as

Electronic copy available at: https://ssrn.com/abstract=3634804
compared to multiple-choice test items created using the 2016 NCSBN® Test. The findings were used to support the hypothesis, or the null defined as follows:

Null Hypothesis: There are no statistically significant differences in the mean scores of the multiple-choice test items developed using the NCSBN® NCJ Model as a framework, as compared to the multiple-choice test items developed without using the framework.

Hypothesis: There are statistically significant differences in the mean scores of the multiple-choice test items developed using the NCSBN® NCJ Model as a framework, as compared to the multiple-choice test items developed without using the framework.

Data Management

The researcher was the only person with administrative access to the examination and the examination results. Participants were restricted to viewing the examination only and were privy to the cohort results. Access to the LMS was closed immediately upon submission of the examination by the participant. The results were accessible to the researcher only and were downloaded to the researcher’s personal computer in the researcher’s home office which is locked when not in use. Both the personal computer and the computer file where the information is stored is password protected. The files were backed up to an external hard drive that is password protected and stored in a locked fireproof safe.

Scope of the Project

The scope of this project was limited to the results of the scores of a convenience sample of prelicensure nursing students on an exam containing 30 test items. Inclusion criteria for the population were defined as any nursing student in the second year of the associate degree nursing program in the approved community college located in the Pacific Northwest. The exact sample size could not be precisely determined as participation was voluntary. The potential available
population consisted of a cohort of 70 students. Exclusion criteria included students in jeopardy of failing their current nursing course as determined by an average grade of less than 80% by the course lead faculty. Data collection took place during the Spring Quarter, May of 2019.

**Limitations of the Project**

The quasi-experimental design of this study lacked the rigor of a true experimental study consisting of an experimental group of students exposed to the independent variable and a control group of students not exposed to the independent variable. A true experimental study was not deemed appropriate for students who were actively enrolled in a nursing program of study due to the potential impact of disrupting the students’ learning (Creswell, 2012). The quasi-experimental design allowed volunteers to participate without impacting their standing in their current courses. All students were provided the same examination, and a comparative analysis of the findings was performed between the two groups of test items.

A large randomized cross-section sample of prelicensure undergraduate students in ADN and BSN programs at community colleges, non-profit, for-profit, public and private universities throughout the country would have provided a more accurate analysis of the test items (Polit & Beck, 2004). The use of a convenience sample of volunteers from only one ADN community college limits the findings of this study. Due to the time limitations of gaining approval from multiple review boards and access to multiple programs across the country a large randomized study was not possible. An invitation for participation was extended to the entire accessible population requesting volunteers for participation in the study. A potential lack of volunteers could have led to a small sample size thereby limiting the results of the findings. If there would have been insufficient volunteers from the original cohort, permission would have been requested to access a similar cohort at a second community college.
Creation of the independent variables which were the multiple-choice test items by only one researcher also limited the findings of this study (Norwood, 2000). Incorporating the participation of multiple nursing faculty from a variety of undergraduate nursing programs would have provided a more accurate evaluation of the ability of the nurse educators to follow the specified framework. Nursing faculty volunteers were not able to be obtained for the duration of this study due to the overload status of the nursing faculty in the ADN program at the participating community college.

The scores on this exam did not have any academic ramifications for the students. Participants may not have taken the examination seriously and may not have applied their best efforts. The need to practice NCLEX-RN® style questions to pass the licensure examination was emphasized. The information flyer indicated that a test review would be conducted by the proctor after the examination and would include the rationales for the correct and incorrect answer choices on each test item. Students may have felt this information was beneficial in preparing for the NCLEX-RN® examination.

Participation of subject matter experts may have been insufficient to assist in the development of high-quality test items that presented realistic client situations. In discussing this project with nurse educators, practicing nurse clinicians, and physicians, content expert volunteers were recruited to review the test items for clinical accuracy and fidelity. Several faculty indicated they have written items for NCSBN®.

The nursing faculty proctor could have provided students with information that guided their answer selections during the exam. A written script of specific instructions was provided, and the proctor was asked not to deviate from the script (Appendix I). Students were informed no questions would be allowed after the examination began.
Technical malfunctions or inaccessibility of the computer software could have occurred. Assistance from information technology service technicians was available during the examination. Backup paper copies of the exam with optical mark recognition answer sheets were available for use in case the system was down.

**Delimitations of the Project**

The limitations of time and resources to develop complex technological enhanced test items, and the lack of access to large, diverse populations of undergraduate RN students in both ADN and BSN programs in a variety of public and private institutions was a limitation to the generalizability of this research. The data for this study was obtained from ADN undergraduate nursing students in their second-year of study in a concept-based curriculum from one community college program only. The nursing program selected for this study is conducted in quarters as opposed to semesters. The results may not be considered generalizable to BSN programs conducted in semesters, or programs based on the medical-model instead of the concept-based model.

Students requiring special accommodations for test anxiety were excluded as all participants were required to take the exam in the same room during the scheduled date and time. However, the scope of this study was not to determine the student’s score on the examination due to either their level of knowledge or the impact of test anxiety on their ability to perform, but rather how well the independent variables, the multiple-choice test items, functioned in eliciting the necessary data for analysis to determine the student’s cognitive operational level on the clinical judgment continuum. The student’s level of understanding of the content of the courses in the curriculum was not evaluated by this study.
The only type of test item examined in this study were the multiple-choice test items. Alternate format items such as drag-and-drop or fill-in the blank items were not tested. Creation of the multiple-choice test items by only one nurse educator, the researcher, was recognized as a potential limitation. The researcher obtained peer review of all items by SME, as well as the interpretation of the analysis of the results in hopes of eliminating bias.

**Feasibility of the Project**

**Costs.** The cost to conduct this project was minimal. Transportation costs to and from the research facility approximately 4 times cost approximately $30.00 dollars. Printing of the recruitment flyers was $5.00. Printing of the examination and purchase of the scantrons was approximately $20.00. These were available to be used if the online function of the LMS was not available. Printing of the certificates of participation for the volunteers on high-quality paper cost $25.00. Gift cards for the subject matter experts and the proctor totaled $150.00. The total cost to conduct this project was $230.00.

**Permission to access the research site.** Preliminary permission was granted by the Dean of the nursing program to conduct this project with the second-year associate degree nursing cohort (Appendix J). The community college required proof of approval from the Institutional Review Board at American Sentinel University prior to granting official approval. The Dean facilitated access to the cohort for the recruiting presentation as well as scheduling of the nursing computer lab to conduct the examination. The faculty member recommended by the Dean proctored the examination.

**Resources.** A free Canvas® course was obtained to deliver the examination. The Dean granted permission to use the nursing programs computer lab to administer the examination. Five subject matter experts reviewed the test items. The Dean at another local community
college granted use of a Scantron® machine if optical scanners were needed, however, this was not necessary.

**Time.** The Dean and the lead faculty of the quarter in which the project was conducted determined the time and date that was most convenient for the students to participate in this study. Every attempt was made to limit the impact of this study on the students’ current coursework. Students were granted one hour to complete the 30-item examination. Set-up time took approximately 15 minutes and no technological difficulties were encountered. One hour was allotted for item review and discussion. Participants were not required to stay for the review session.

**Theoretical Framework**

The theoretical framework selected to guide this study was Orlando’s Nursing Process Theory (Schmieding, 2002). The goal of this research was to evaluate the effectiveness of the NCSBN® NCJ Task Model as a framework to use to develop multiple-choice test items that accurately measure the clinical judgment ability of prelicensure nursing students. The cognitive activities prompted by the cognitive operations of the NCSBN® NCJ Task Model provided the basis for the contextual factors utilized in the creation of the multiple-choice questions which were the independent variables of this study. The resulting constructs formed from the cognitive operations elicited by the multiple-choice questions were the student’s answer selections which were the dependent variables.

The NCSBN® NCJ Assessment Model was used to analyze the answer choices. Results were categorized using the layers of the NCSBN® NCJ Assessment model to determine the level of clinical judgment ability of the student (Dickison et al., 2016). The Nursing Process Theory (NPT) was selected because it systematically identifies the constructs of the practice of nursing.
and links these constructs to the cognitive operations essential to creating clinical judgments represented in the NCSBN® NCJ Task Model and the layers in the NCSBN® NCJ Assessment Model. A model depicting the relationships of these constructs was provided in Appendix K.

Orlando developed the NPT in 1958. Orlando was one of the first to emphasize the importance of the nurse as an independent thinker responsible for meeting the client’s needs to facilitate optimal health outcomes. She emphasized the importance of the nurse-client relationship in which the nurse verifies the assessment findings with the client prior to formulating a decision on the course of action to take to meet the client’s needs (Schmieding, 2002). Orlando’s original theory consisted of the three steps of assessment, planning, and evaluation. The NPT has been revised over time and today includes the five-step process of assessment, diagnosis, planning, implementation and evaluation (Doenges & Moorhouse, 2013). The major assumptions of the NPT relevant to the exploration of the measure of clinical judgment proposed in this research study include (Schmieding, 2002, p. 402):

1. Nursing is a distinct profession that requires nurses to act independently of other disciplines to meet the client’s needs
2. Clients are individuals with unique needs that can be communicated verbally and nonverbally
3. Nurses gain knowledge from interaction during the nurse-client encounter
4. Nurses are responsible for validating assessment findings with the client and implementing competent nursing actions to meet the needs of the client

The NPT is relevant to the discipline of nursing today. The NPT has been used as a conceptual framework to investigate issues identified in nursing practice, as well as in nursing
educational settings. The validity of the NPT as a reliable conceptual framework has been established by numerous research studies over the past 50-years (Schmieding, 2002).

Shabel (1989) used the NPT to evaluate if registered nurses use the steps of the nursing process in their day-to-day practice. Shabel used a questionnaire to obtain information about the nursing process related to the practice of RNs working at 17 hospitals. The results from the 322 surveys returned indicated that 94% of the RNs considered the systematic planning of nursing care is a professional responsibility; 88% indicated the nursing process is essential to quality nursing practice; and 86% believed the use of the nursing process improves the quality of client care provided (p. 77).

Potter and Bockenhauer (2000) conducted a pilot study to determine the effectiveness of the NPT in improving nurse-client interactions at a large psychiatric facility. The researchers hypothesized the commitment by the nursing staff to follow the guidelines of Orlando’s NPT would help the nurses to achieve excellence in client interactions and thereby improve the client outcomes. The clients’ levels of distress were measured using the Bockenhauer/Potter Scale of Immediate Distress. The independent variable was the use of the NPT during client interaction. Interactions were documented based on the use of the NPT, or interactions not using the NPT. The findings of the study indicated a reduction in client distress levels, improvement in client outcomes, and self-reported ”greater sense of purpose and direction” by the RNs when the NPT was used to guide the client-nurse interactions (p. 20).

The NPT is also a verified conceptual framework for use in nursing education programs. The steps of NPT are identified by the NCSBN® as fundamental processes and are integrated throughout the NCLEX-RN® test plan (2016). McEwen and Brown (2002) surveyed 300 undergraduate nursing programs to determine the conceptual frameworks used in the curricula.
The results revealed the NPT was the most widely used conceptual framework (55%) for the nursing curricula throughout the undergraduate nursing programs surveyed.

The NPT guides the legal scope of practice for nursing. The steps of the NPT serve as foundational legal definitions of the nursing scope of practice by state boards of nursing and are outlined in the American Nurses Association (ANA) Standards of Professional Nursing Practice (ANA, 2015b; Doegnes & Moorhouse, 2013, p. 4-7). Johnson (2013) examined the results of litigation claims of allegations of negligence brought against RNs practicing in perinatal units in Florida. Johnson used a qualitative summative approach to categorize and rank the allegations according to number of occurrences per each step in the nursing process. The implementation step was the most prevalent category \( n = 160 \), followed by the diagnostic \( n = 150 \) and planning \( n = 125 \) categories. Johnson recommended nurse leaders develop guidelines to promote adherence to the nursing process in order to be in compliance with the Florida Nurse Practice Act.

The NCSBN® defines the nursing process as “a scientific, clinical reasoning approach to client care that includes assessment, analysis, planning, implementation and evaluation” (2016, p. 5). Alfaro-LeFevre (2017) emphasizes the systematic organization of the nursing process as the “foundation for clinical judgment” (p. 71). Doenges and Moorhouse (2013) declare the nursing process is “the basis for all nursing actions and is the essence of nursing” (p. 8).

Nurse educators are responsible for creating learning experiences in prelicensure nursing programs that promote the use of the steps in the nursing process by nursing students and the development of critical thinking skills that lead to competent clinical judgments (NLN, 2013). The purpose of this research study was to evaluate the effectiveness of the NCSBN® NCJ Task Model as a framework to use to create multiple-choice test item that accurately assess the clinical
judgment abilities of prelicensure RN students. The NCSBN® NCJ Task Model consists of five
cognitive operations which discussed in the following narrative to provide an explanation of the
relationship to the steps in the nursing process. These five sequential cognitive operations are
the decision-making actions essential in clinical judgment (Dickison et al., 2016). However,
these cognitive operations are general processes that can be applied to all types of disciplines or
any situation requiring critical thinking to determine the most appropriate outcome. The NPT
provided a systematic framework that linked these cognitive operations to the discipline of
nursing.

Step one of the NPT is assessment. Nursing assessments consist of information obtained
from subjective and objective data gathering. The subjective assessment is derived from talking
with the client, the client’s family members or significant other, the members of the
interdisciplinary health care team and review of the client’s medical record. The objective data
is obtained by performing a physical assessment and reviewing the findings of diagnostic tests
and procedures (Doenges & Moorhouse, 2013). Assessment requires the cognitive activities of
recognizing and recalling (Anderson et al., 2001). Assessment closely aligns with the cognitive
operation of cue recognition in which information is recognized as meaningful. Muntean et al.
(2015) identified cue recognition as the initial step in the clinical judgment process.

Step two of the NPT is diagnosis. In this step, the nurse analyzes the assessment data to
differentiate normal findings from abnormal findings to determine the client’s healthcare needs
(Doenges & Moorhouse, 2013). This step requires the cognitive activities of understanding,
interpreting, classifying, analyzing, differentiating, and organizing (Anderson et al., 2001). The
diagnosis step of the NPT aligns with generating hypotheses, the second cognitive operation task
of the conceptual model of nursing clinical judgment (Muntean et al., 2015). To generate
accurate hypotheses, the RN student must accurately interpret the relevancy of assessment information in accordance with nursing knowledge and classify the information into groups to determine the client’s healthcare needs (Hooper, 2017, November).

Step three of the NPT is planning. Planning is a two-tiered process. First, the RN student identifies expected outcomes for the client. Expected outcomes are the desired results the client hopes to achieve. These are sometimes referred to as long-term goals and indicate the progression of the client towards optimum health. Second, the RN student defines the steps, or short-term goals, that are necessary to attain the expected outcomes (Doenges & Moorhouse, 2013). The ability to differentiate the classifications of groups and to structure the order of the steps in the desired goals is essential to developing appropriate client outcomes (Anderson et al., 2001). The planning step correlates to the third cognitive operational task of prioritizing the hypotheses (Muntean et al., 2015). This step requires the RN to organize the hypotheses into a list of client’s needs prioritized to address the most urgent needs first (Hooper, 2017, November).

Step four of the NPT is implementation. This step consists of nursing interventions that will be performed to meet the needs of the client identified in steps one through three. These nursing actions must be evidence-based and adhere to current standards of nursing practice (Doenges & Moorhouse, 2013). The cognitive activity of application is required in this step. The RN student must decide which nursing actions to execute to safely care for the client (Anderson et al., 2001). This step directly aligns with the fourth step of the cognitive operation task which is the take action step (Muntean et al., 2015).

The final step in the NPT, step five, is evaluation. In this step, the RN student reassesses the client to determine if the nursing strategies are producing the desired therapeutic effects. The RN student must critique the re-assessment findings to determine if the nursing interventions are
producing the therapeutic effects intended, if the client is progressing towards the desired outcomes identified in step three, thereby indicating an improvement in the client’s health status. This is a crucial step in maintaining client safety and promoting positive client outcomes. If the reassessment findings indicate a lack of therapeutic response, the RN student must then decide how to edit the plan of care to incorporate new nursing strategies that will prevent client decline (Doenges & Moorhouse, 2013). This step requires the higher-order cognitive activity of evaluation in which the student must judge the consequences of the actions implemented in step four (Anderson et al., 2001). This step directly correlates with the final step of evaluating outcomes identified by Muntean et al. (2015) in their conceptual model of nursing clinical judgment. As in step one, cue recognition is essential in identifying the client’s response to the nursing actions.

The cognitive activities identified in the cognitive operational tasks required to develop clinical judgments indicate each step of the NPT is supported by the decision-making actions. The steps in the NPT allowed for the incorporation of the nursing perspective into the development of the multiple-choice test items using the NCSBN® NCJ Task model thereby linking these concepts to nursing practice (Dickison et al., 2016). The nursing process is a continuous repetitive cycle that begins with the first encounter of the client, recurs with every subsequent encounter, and ends with the last encounter with the client. During each client-nurse encounter, the nurse is responsible for recognizing assessment cues that indicate client distress, prioritizing the nursing diagnoses generated from the analysis of the assessment data, developing goals, formulating competent interventions based on current standards of care, and then evaluating the client outcomes.
Each of these client-nurse encounters was viewed as nodes on the health continuum (Appendix L). A nursing clinical judgment must be formulated during each encounter because every encounter ends with an outcome that is significant to the client’s progression towards optimal health, or the lack thereof. This process is influenced by the nurse’s level of knowledge, experiences, and the amount of interaction time with the client. Failure to recognize cues of client distress will prevent the RN from working through the steps of the NPT and can have devastating effects on the client’s health status.

The quality of the multiple-choice test items developed for this study was pivotal in accurately evaluating the RN student’s ability to formulate safe clinical judgments based on evidence-based practice standards. The NCSBN® NCJ Assessment Model was used to analyze the student answer selections to categorize the student’s clinical judgment level by using the layers defined by this model (Dickison et al., 2016). The NPT provided the framework that linked the concepts of the clinical judgment operational tasks to facilitate the method of measurement defined by the NCSBN® NCJ Assessment model (Dickison et al., 2016). The NPT was chosen as the conceptual framework for this research study because this conceptual theory:

1. Bridged the gap between the iterative decision-making processes of the NCSBN® NCJ Assessment Model and the nursing scope of practice standards.

2. Established inter-rated links between the cognitive operations identified on the NCSBN® NCJ Task Model to the NPT to guide the development of the multiple-choice test items.

3. Facilitated the analysis of the findings with the NCSBN® NCJ Assessment Model within the scope of the discipline of nursing.
4. Provided a common language to facilitate communication of findings within the nursing body of knowledge.

The NPT facilitated the design, implementation, and analysis of the variables of this project from a nursing perspective. While evaluating other theories for use for this project, such as Knowles Adult Learning Theory (1980), Bloom’s Taxonomy of Cognitive Levels (1956), Miller’s Information Processing Theory (1956), Benner’s Intuitive-Humanistic Model (Brykczynski, 2002), and Kolb’s Experimental Learning Cycle Theory (Littlewood & Demian, 2015), the importance of utilizing a theory that facilitated the evaluation of the concept of clinical judgment within the standards of nursing practice was recognized. The concepts of the NPT aligned this study specifically to nursing education and nursing practice. The major concepts and theoretical assertions of the NPT directly aligned with the definition of clinical judgment by the NCSBN®, as well as the cognitive operational definitions and subcomponents of the NCSBN® NCJ Model. All of the other theories considered did address individual subcomponents of the cognitive operational levels of clinical judgment as defined by the NCSBN®, however, each was lacking in addressing all of the subcomponents of the NCSBN® NCJ Model in assessing clinical judgment using multiple-choice examine items in the didactic educational setting. For example, Benner’s Intuitive-Humanistic Model (Brykczynski, 2002) and Kolb’s Experimental Learning Cycle Theory (Littlewood & Demian, 2015) was excellent in direct observational educational activities such as simulation, clinicals, and skills labs where specific tasks must be performed allowing the nurse educator to assess the psychomotor functions of the student as well as ask open-ended questions to probe the student’s critical thinking process. With written examinations in the theoretical classroom setting this is not
possible. The Miller’s Information Processing Theory (1956) was excellent for generalized critical thinking components for any topic but lacked direct ties to nursing practice.

Definition of Terms

*Ability:* The conceptual definition of ability as it relates to this study was defined as the competence, proficiency, or aptitude in performing the cognitive processes to make a decision (Merriam-Webster's online dictionary, n.d.). Ability was measured by analyzing the student’s selections of the answer choices on the multiple-choice questions using the cognitive operational tasks of the NCSBN® NCJ Task Model (Dickison et al., 2016).

*Analyze cues:* “The candidate interprets cues from their existing knowledge base and nursing perspective, evaluate cues in terms of relevancy, importance and interrelationship among other cues, organize cues in the mental representation of the scenario (e.g., organize cues in clusters), and then develops a group of probable client needs/concerns and problems” (Hooper, 2017, p. 71).

*Assessing:* The conceptual definition of assessing was defined as the act of judging or evaluating something (Merriam-Webster's online dictionary, n.d.). Assessing was operationalized in this study by evaluating the level of clinical judgment achieved by the student from the student’s answer selections on the multiple-choice questions as measured by the NCSBN® NCJ Assessment Model (Dickison et al., 2016).

*Candidate:* A graduate of a registered nursing program of study that has been granted permission to take the NCLEX-RN® examination (NCSBN, 2015b).

*Clinical judgment:* The conceptual definition of clinical judgment as it pertains to this project was defined by the NCSBN® as “an iterative decision-making process that uses nursing knowledge to observe and assess presenting situations, identify a prioritized client concern, and
generate the best possible evidence-based solutions in order to deliver safe client care” (Woo, 2016, October, slide 5). The measurement of clinical judgment was operationalized by creating the multiple-choice questions to present client assessment data in the form of small vignettes, followed by questions requiring the RN student to make the safest decision regarding the client’s nursing care.

Clinical judgment element descriptors: Cognitive operations used to construct clinical judgments: a) recognize cues, b) analyze cues, c) prioritize hypotheses, d) generate solutions, e) take actions, and f) evaluate outcomes (Hooper, 2017, p. 71; Muntean et al., 2015).

Cognitive activities: Cognitive is defined by Merriam-Webster (2019) as “involving conscious intellectual activity.” Cognitive activity was measured in this study using Anderson and Krathwohl’s revised Bloom’s taxonomy cognitive domains (Anderson et al., 2001).

Cognitive operational tasks: The cognitive operational tasks are the cognitive activities required in the process of clinical judgment as identified by Dickison et al. (2016). These were measured using the NCSBN® NCJ Task Model.

Effectiveness: The conceptual definition of effectiveness is producing the desired effect or outcome (Merriam-Webster, 2019). Effectiveness will be measured by analyzing the student’s answers to the multiple-choice questions to determine the cognitive operational level on the NCSBN® NCJ Task Model (Dickison et al., 2016).

Framework: The conceptual definition of framework was defined as a structure that provides the basis for the development of ideas (Merriam-Webster, 2019). Framework was operationalized for this study by the utilization of the NCSBN® NCJ Task Model as the framework that was followed to create the multiple-choice test items that served as the independent variables (Dickison et al., 2016).
Level of cognitive operation: The levels of cognitive operations were the layers of the construct factors of cognitive operations as identified by Dickison et al. (2016). These were measured using the NCSBN® NCJ Assessment Model.

Multiple-choice test item: The conceptual definition of multiple-choice was a question on an examination that consisted of several choices from which the answer must be chosen (Merriam-Webster's online dictionary, n.d.). The multiple-choice test items were the independent variables in this study and were developed following the guidelines from the NCSBN® NCJ Task Model (Dickison et al., 2016).

NCSBN® Nursing Clinical Judgment (NCJ) Assessment Model: The multilayer model developed by the NCSBN® consisting of the “cognitive operations and contextual factors” processes required to formulate clinical judgments (Dickison et al., 2016, p.4). This model was used to categorize the level of clinical judgment achieved by the test takers of this study.

NCSBN® Nursing Clinical Judgment (NCJ) Task Model: The model developed by the NCSBN® to “facilitate the development of highly structured items that elicit responses and generate data from test takers in a consistent manner” (Dickison et al., 2016, p. 7). This model was used as the framework for the development of the multiple-choice test items for this study.

Nurse educator and nursing faculty: For the purposes of this study nurse educator and nursing faculty were used interchangeably and were defined as a registered nurse with a master’s degree in nursing who teaches in an undergraduate nursing program (NCSBN, 2008).

Prelicensure: Prelicense is defined by Merriam-Webster (2019) as “occurring prior to the issuance of a license.” For the purposes of this study, the operational definition of prelicensure is prior to obtaining a license to practice as a registered nurse in the United States. Successful completion of the NCLEX-RN® examination is required to obtain an RN license. To be eligible
to take the NCLEX-RN® examination candidates must complete an approved undergraduate nursing program of study (NCSBN, 2017). For this project, the nursing program of study was an associate degree nursing program.

*Prelicensure RN students:* A prelicensure RN student was identified for the purposes of this study as a student who is enrolled and in good standing in an associate degree of nursing program (NCSBN, 2017).

*Registered nurse:* A registered nurse is defined as “an individual who has graduated from a state-approved school of nursing, passed the NCLEX-RN Examination and is licensed by a state board of nursing to provide patient care” by the NCSBN® (2017).

*Vignette:* The conceptual definition of vignette is a short description of something (Merriam-Webster, 2019.). Vignettes were used to provide the client information for the framework test items in this study.

**Summary**

The ability to formulate safe competent clinical judgments is a requirement of all RNs who care for clients (Jessee, 2018; Kavanagh & Szweda, 2017; Makary & Daniel, 2016). Nursing faculty in prelicensure nursing programs are responsible for assisting students to develop the ability to provide safe client care as a novice nurse (NLN, 2013; Purling & King, 2012). Accurately assessing every student as the students’ progress through the nursing program is essential to verify all students are grasping the theoretical content and can apply the content to a client care situation. Nurse educators use theoretical examinations to assess large cohorts of nursing students (Masters et al., 2001; Tarrant et al., 2006).

The review of the literature revealed research was lacking in evidence-based assessment tools to develop test items that accurately assess clinical judgment (Bristol et al., 2018; Clifton &
Schriner, 2010; DePew, 2001; Johnson, 1999; Schroeder, 2007; White & Heitzler, 2018). This gap supported the pursuit of the research question proposed in this project: can multiple-choice questions accurately assess the clinical judgment abilities of prelicensure registered nursing students? This research was significant to nursing education because the findings were evidence-based and provided insight into a process for developing multiple-choice test items that were structured according to the new NCSBN® item format purposed for the upcoming revisions to the NCLEX-RN® examination (NCSBN, 2017, Fall).

This study added to current nursing knowledge and presented an evidence-based method for developing multiple-choice test items and analyzing the results that can be used with large cohorts to accurately measure clinical judgment. By analyzing individual student performance, nurse educators can identify struggling students and assist the students to identify study strategies that will facilitate the development of the higher-order critical thinking skills necessary to apply the theoretical knowledge to a clinical client care situation. These strategies benefit society in general by improving the ability of novice RNs to safely care for clients.

Section I has explained the background and significance of the basis for this research project. Section II describes the methods that were used to conduct this quasi-experimental study, including the sample, the setting, instrumentation and data collection plans. Section III provides the results of the findings from the statistical analysis with recommendations for nursing education and future research.
SECTION II: METHODS

Introduction

Evidence-based assessment methods are needed to evaluate registered nursing (RN) student’s understanding of the higher-order cognitive constructs of clinical judgment in undergraduate nursing education programs. Bristol et al. (2018) found there was a lack of evidence in the current nursing literature on the best practices for test item development, administration, and analysis. The purpose of this study was to evaluate the effectiveness of multiple-choice test items in accurately assessing the clinical judgment abilities of prelicensure RN students.

The National Council of State Boards of Nursing is revising the Nursing Council Licensing Examaintion for Registered Nurses to provide a more accurate measurement of the concept of clinical judgment. The NCSBN® has implemented the Next Generation Project to improve the test items on the NCLEX-RN® examination with the goal of accurately assessing the RN candidates’ ability to develop competent clinical judgments which are essential to safe patient care (NCSBN, 2017, Fall). Evidence-based assessment methods are needed to accurately assess the clinical judgment abilities of large cohorts of RN students. By performing this study, the researcher sought to answer the question: Can the NCSBN® NCJ Task Model be used as a framework to develop multiple-choice test items that accurately measure the clinical judgment abilities of pre-licensure RN students?

Project Design

A quantitative design was chosen to facilitate measurement of the outcome of clinical judgment and demonstrate the impact of the independent variable, the multiple-choice questions, on the dependent variables, the students answer choices. The conceptual definition of clinical
judgment, which was the outcome of this project, was defined by the NCSBN® as “an iterative decision-making process that uses nursing knowledge to observe and assess presenting situations, identify a prioritized client concern, and generate the best possible evidence-based solutions in order to deliver safe client care” (Woo, 2016, October, slide 5). To operationalize the outcome of clinical judgment the NCSBN® Nursing Clinical Judgment Task Model was used as a framework to create multiple-choice questions to serve as the independent variables (Dickison et al., 2016).

The test items developed using the NCJ Task Model presented realistic client assessment data in the form of small vignettes, followed by serial multiple-choice questions that required the RN student to analyze the information and form a clinical judgment regarding the client’s nursing care (Appendix D). The multiple-choice test items were developed using the cognitive operational categories defined on the NCJ Task Model of cue recognition, cue analysis, hypothesis development and prioritization, generation of solutions, and evaluation of outcomes (Dickison et al., 2016; Hooper, 2017, November). These cognitive operational functions were the basis for both the NCSBN® NCJ Task Model and the NCSBN® NCJ Assessment Model. The outcome of clinical judgment was measured using the clinical judgment element descriptors defined by the NCSBN® Nursing Clinical Judgment Assessment Model to analyze the dependent variables which are the participants’ answer selections (Dickison et al., 2016). The numerical value assigned to each of the cognitive operational levels were used to quantify the measurement results for analysis using descriptive and inferential statistics.

A quasi-experimental approach was selected to allow RN students to participate voluntarily without impacting their course studies. A quasi-experimental study lacks a control group and randomization but has less potential for disrupting student learning (Creswell, 2012).
All students received the same test items and the same information regarding the rationales for the correct answers. By developing an experimental group of test items using the new framework (Appendix D), and a comparative group of test items (Appendix F) using current guidelines, a comparative analysis of the findings could be performed to determine if the new process of test item development was feasible and transferable to practice in nursing education (Norwood, 2000).

Sample and Setting

The characteristics of the population, methods used to select the sample population, and the process of determining the size of the sample are discussed in this section. Explanation of the criteria for inclusion and exclusion are delineated, as well as the strengths and limitations of the sampling strategy.

Characteristics of Target and Accessible Population

The target population for this study was the prelicensure RN students in undergraduate RN programs throughout the United States. Prelicensure RN programs consist of Associate Degree Nursing programs (ADN) and Baccalaureate Degree Nursing programs (BSN). RN-to-BSN programs were not considered part of this target population as students in these types of programs have completed the NCLEX-RN® examination and have an RN license.

The accessible population for this study was the RN students in the second-year of an ADN program in a community college in the Pacific Northwest region of the United States. The setting for the study was a community college in the Pacific Northwest. The community college offers an Associates in Nursing Direct Transfer Agreement/Major Related Program (DTA/MRP) degree. Students complete 135 credits at the community college and are then eligible to take the
NCLEX-RN® examination and to transfer to a university to complete a bachelors in nursing (BSN).

The cohort consisted of 70 students, the majority of whom were Caucasian females between the ages of 22 to 35. Demographics were collected by requesting the participants to provide their age, gender, race, overall grade point average (GPA) range, previously completed college degree or certificate programs, and work experience in a healthcare environment. All students had completed fundamentals and medical/surgical clinical rotations and had an overall GPA of 2.2 or above as this was required to remain in this nursing program. The program design was a concept-based curriculum which aligned well with the concept-based format of the test items in this study.

If time did not allow the completion of this study before the end of the Spring 2019 quarter, permission would have been requested to complete the study over the summer semester with a group of BSN students in their junior year of study at a private university in the Pacific Northwest. The BSN cohort consisted of 50 students, the majority of whom are also Caucasian females between the ages of 22 to 35. The Dean of the BSN program had granted preliminary permission to conduct the study at the university, pending IRB approval (Appendix J). Every attempt however was made to complete the study at the community college due to the larger cohort size in hopes of obtaining a sufficient number of volunteers.

**Proposed Sampling Strategy with Strengths and Limitations**

The sampling plan for this quasi-experimental study was a nonprobability sample due to the limitations of subject availability and preferential selection of second-year ADN students (Tappen, 2016). A large randomized cross-section sample of associate degree students from undergraduate nursing programs throughout the country would have provided a more accurate
analysis of the test items. Upon advisement of faculty, a purposive sample from one prelicensure program was selected over a large random sample from multiple prelicensure programs due to the time constraints required to complete the study and the necessity to complete the approval process for multiple Institutional Review Boards (IRB).

While the use of a convenience sample from only one community college limited the findings of this study, a nonprobability purposive sample was advantageous since the goal was to explore the effectiveness of a new process for developing multiple-choice test items to evaluate clinical judgment in prelicensure RN programs. Polit and Beck (2004) indicate that a purposive sample can provide an effective evaluation of a newly developed instrument. True experimental studies consisting of an experimental group and a control group provide a more accurate assessment of the impact of the independent variable on the dependent variable. This quasi-experimental study lacked a control group and randomization. However, a true experimental study was not feasible to conduct with a group of students actively pursuing a degree due to the potential for disrupting their learning (Creswell, 2012). A degree of randomization was accomplished during the analysis phase of this study by performing a random selection of the participants’ results using a computer-generated list. Analysis using stratified sampling by clustering the participants based on the demographic categories were also performed. “The goal in each of these methods was to achieve randomness and avoid any systematic procedure that might introduce bias” (Tappen, 2016, p. 121).

Potential limitations of the accessible population for this study involved the type of nursing program. This ADN program was a concept-based program administered in six quarters over a two-year period. Programs conducted using the medical-model over a 16-week semester may have increased the length of exposure to clients in the clinical setting, thereby improving the
development of students’ clinical judgment abilities. Future studies will need to be conducted to determine if the findings of this study are relevant to programs based on the medical-model plan of study and programs structured in semesters instead of quarters.

Projected Sample Size

**Power analysis using effect size.** The sample size necessary to determine the statistical significance between the two groups of independent variables which consisted of two groups of test items was estimated at 42. The estimated sample size was derived by conducting a priori power analysis using the G*Power software with the following settings: (a) two tail t-test, (b) effect size 0.8, (c) alpha 0.05, and (d) power beta of 0.95 (Faul, Erdfelder, Buchner, & Lang, 2009). Setting the alpha at a level of 0.05 or lower reduced the risk of making a Type I error (Polit & Beck 2004, p. 508).

**Power analysis using confidence intervals.** The Raosoft® sample size calculator (2004) was used to estimate the necessary sample size using the following confidence level settings:

- 99% (1% margin of error) sample size = 64
- 95% (5% margin of error) sample size = 60
- 90% (10% margin of error) sample size = 56

The size of the accessible population was 70 students who are representative of the target population. The majority of the cohort met the inclusion criteria and volunteered to participate in the study which ensured a sample size large enough to minimize the possibility of a Type I or Type II error.
**Inclusion Criteria**

All students in good academic standing in the second-year ADN medical/surgical class were considered eligible to participate in the study. Good academic standing was defined as a minimum grade point average of 2.2, or 80% of the total course points, which is required to continue in the nursing program. Participants were encouraged to complete the demographic section and all 30-items on the examination, however, all responses were included in the analysis of the findings of the study.

**Exclusion Criteria**

Any second-year ADN student with a current course average less than 80% were not allowed to participate. Every attempt was made to minimize the impact of this study on the students’ ability to successfully complete the current nursing courses in which they were enrolled. Students who were in jeopardy of failing their current courses would have been deemed ineligible for inclusion. All participants were required to take the examination together at the same time in the nursing department computer classroom. Any student who requested to take the examination at a different time or in a different location would have been excluded from the study.

**Contextual Factors of the Potential Impact on the Outcome**

A lack of volunteers would have led to a small sample size which would limit the generalizability of the findings and increased the probability of a Type II error (Polit & Beck, 2004). If the number of volunteers at the initial college were less than 42, the researcher would have requested permission of the Chair to access the ADN cohort at a second community college in the area that was also a concept-based program. The size of the second-year ADN cohort at the second college was 50 providing a total potential population of 120. Preliminary permission
to conduct the study was obtained from the Dean at both colleges pending IRB approval (Appendix J). After the study was approved by the American Sentinel University IRB (Appendix M) permission from the college IRB was obtained (Appendix N).

Students could have been hesitant to volunteer fearing embarrassment if the results of their examination scores were low. During the recruitment phase, students were assured that the test was strictly voluntary, and the scores would not impact their course grade. The students may have feared they were not prepared for the subject matter addressed by the test items. Students were reassured of exposure to the concepts during previous nursing courses and an explanation provided of how anonymity of the results would be maintained in hopes of alleviating these fears.

Students who volunteer may not have taken the examination seriously since the scores would not impact their course grades. Students at this level are usually eager to practice NCLEX-RN® style questions as they will be taking the licensure exam in the upcoming months. The information flyer (Appendix O) indicated that a test review would be conducted by the proctor after the examination that included the rationales for the correct and incorrect answer choices on each test item. Students may have felt this information was beneficial in preparing for the NCLEX-RN® examination.

RN students are usually very pressed for spare time. The recruitment process was timed with the faculty contact person to be presented at the least perceived stressful time during the quarter. Specific attention to assignment due dates and examinations dates were taken into consideration. In addition, the number of items on the examination was limited to 30 to allow the students to complete the examination in 60 minutes or less. A larger sample of items would have provided more data to support the validity of the level of measurement by providing more
data for analysis of internal consistency. However, the increase in the amount of time necessary to complete a longer examination would have deterred students’ willingness to participate in the study.

Repeatability of the process used to create the multiple-choice questions was an essential outcome of this study as the results will hopefully provide evidence supporting the item-development framework for use by nursing faculty. The purpose of this project was to explore the effectiveness of using the NCSBN® Nursing Clinical Judgment Task Model to develop multiple-choice test items that provide an accurate assessment of the ADN student’s clinical judgment abilities (Dickison et al., 2016). The researcher attempted to prevent bias during item development by eliciting input from subject matter experts to review the items for accuracy. Explicit, detailed documentation of each step in the item development process was maintained by the researcher. Clear delineation of the definition of terms, project design, method of item-development, examination procedures, and interpretation of the statistical analysis of the results are provided for ease of replication.

**Instrumentation**

The NCLEX® Examination Committee members commissioned research in 2014 to determine the best evidence-based processes to ensure the NCLEX® RN licensing examination accurately measures safe clinical judgment abilities of the NCLEX-RN® examination candidates. This research project is known as the Next Generation Project (NCSBN, 2017, Fall). After extensive exploration, “a multilayer assessment model” was created to “establish a psychometric representation of nursing clinical judgment” called the NCSBN® Nursing Clinical Judgment (NCJ) Task Model (Appendix B). The multilayer design of this Task Model facilitates the design of test items that measure “cognitive subcomponents of NCJ.” The NCSBN® NCJ
Assessment Model (Appendix A) was created to analyze the candidates answer selections to “identify at which process nurses may make errors” (Dickison et al., 2016, pp. 3-4).

Prototype test items were constructed and added as a research section to the NCLEX-RN® examination in July 2017 to test the reliability and validity of the items in measuring the cognitive constructs of clinical judgment (Grossenbacher & Kappel, 2018, Winter). Preliminary test item analysis results indicated the items constructed following the NCSBN® NCJ Task Model accurately assess the cognitive constructs of clinical judgment as defined by the NCSBN® NCJ Assessment Model (Dickison & Woo, 2016, July 25; Grossenbacher & Kappel, 2018, Winter; Muntean et al., 2015, April 10; NCSBN, 2018, Winter; P. Dickison, personal communication, March 23, 2018). Testing of the NCJ prototype items is ongoing. In discussions with Dr. Philip Dickison, Chief Officer of Operations and Examinations for NCSBN®, he stated, “To-date we have collected over 1.5 million data points from the Special Research Section of the NCLEX-RN® examination that began in July 2017. The data consistently shows the reliability of the new item types in accurately assessing clinical judgment that is equal to or better than the item types on the current NCLEX-RN examination.” He went on to discuss the viability of the Next Gen Project, and that no information obtained thus far indicates changes are needed in the NCJ model. Full rollout of the new NCLEX-RN® examination is estimated to take place in 2023. Publication of the findings from the prototype testing is expected to be made available during the Fall of 2019. (P. Dickison, personal communication, February 26, 2019).

The NCSBN® NCJ Task Model was selected as the framework for the multiple-choice test item development because this framework was used by the NCSBN® to create the new NCLEX-RN® test items. The NCSBN® NCJ Assessment Model was used to assess the
cognitive operational levels of clinical judgment. Permission was obtained from the NCSBN® to use the NCSBN® NCJ Models (Appendix P). Both the NCSBN® NCJ Task Model and the NCSBN® NCJ Assessment Model were validated by the NCSBN® as reliable tools for test item construction and analysis (Betts, 2017, November; Dickison et al., 2016; NCSBN, Winter, 2018; P. Dickison, personal communication, March 23, 2018).

Construct validity of the NCSBN® Nursing Clinical Judgment Task Model was established through the results of the initial and subsequent studies performed by Dickison et al. (2016) proving the questions developed using the Task Model provided the information necessary to measure cognitive constructs of clinical judgment as defined by the NCSBN® Nursing Clinical Judgment Assessment Model. The constructs of clinical judgment correlate with the conceptual underpinnings of the Nursing Process Theory by Orlando which was essential to this research project as the NPT provided the basis for the Registered Nurses’ Scope of Practice (American Nurses Association, 2015; Schmieding, 2002).

Reliability of the NCSBN® NCJ Task Model and the NCSBN® NCJ Assessment Model was established using the test-retest method through repeated studies conducted by Dickison et al. (2016). Equivalence reliability of different types of test items was established by using the NCSBN® NCJ Task Model to develop numerous examinations containing a variety of test item types that covered a wide-range of client-care concepts. Internal consistency was proven reliable by creating sibling items for each client-care concept and examining the correlation function of these items using the Assessment Model to guide the analysis of the results related to the outcome of the level of clinical judgment (Dickison et al., 2016). Research is on-going by the NCSBN® and was closely monitored to ensure the results continue to provide evidence of the
reliability and validity of both instruments that were utilized to measure the outcome of clinical judgment for this project.

The examination (Appendices D and F) consisted of 30 multiple-choice test items covering the five physiological concepts of metabolism, perfusion, oxygenation, fluid and electrolytes, and infection. The researcher developed six items per concept. Three of the items were created using the NCSBN® NCJ framework and labeled the framework group (Dickison et al., 2016). Three of the items were developed using the NCSBN® 2016 Test Plan recommendations and labeled the non-framework group (NCSBN, 2015b). The information for the content of the item stems and the answer choices, as well as the rationales for the correct and incorrect responses, was obtained from current, evidence-based textbooks. The references were provided in the rationales section for each item in Appendices D and F.

All 30 items were submitted via email to a panel of subject matter experts (SME) that consisted of three advanced nurse practitioners and three certified nurse educators. Recommendations for improvement of the content of the items and correction of bias language or errors was completed. The revised items were resent to the SME panel for approval. This process was repeated until all items were approved for use by all SMEs. Once the items were approved, the examination was created in the Canvas® open access course using the quiz function.

The physiological concept for item numbers one through six was metabolism. Item numbers one through three were from the framework group. These three items were based on the client information provided in the short vignette in item one. The participants answered questions as the case study unfolded in item two and three. Item number one provided the client case study information containing assessment data of the two physiological concepts of
metabolism and perfusion. Item number one required the participant to recognize the cues provided in the case study, analyze the cues correctly to determine the correct physiological issue, and determine which assessment the nurse should conduct first. This item tested the NCJ cognitive operational layer three on the NCJ Assessment model and correlated to the assessment category of the Nursing Process which was the theoretical framework of this study.

Item number two provided more assessment data. The initial case study information from item 1 was present on the screen if the participants would have needed to access the original information. Item number two required the participant to determine which nursing action to implement. This item tested the NCJ cognitive operational layer two on the NCJ Assessment model and correlated to the implementation category of the Nursing Process.

Item number three asked the participant to evaluate the therapeutic effectiveness of the intervention performed in item number two. The case study information from item one and the additional assessment information from item two was present on the screen in item three if the participants needed to review the information. This item tested the NCJ cognitive operational layer two on the NCJ Assessment model of satisfactorily evaluating outcomes and correlated to the evaluation category of the Nursing Process.

Item numbers four through six were from the non-framework group and were independent items not linked to a case study but pertained to the same physiological concept of metabolism. Item number four was written at Bloom’s taxonomy level of remember and understand and required the participant to recall information specific to assessment of the physiological concept of metabolism. This item addressed the assessment category of the Nursing Process.
Item number five was written at Bloom’s taxonomy level of analyze and apply. Item number five required the participant to interpret the information specific to implementation of a nursing intervention for the physiological concept of metabolism. This item addressed the implementation category of the Nursing Process.

Item number six was written at Bloom’s taxonomy level evaluate. Item number six required the participant to evaluate the therapeutic effectiveness of an intervention specific to the physiological concept of metabolism. This item addressed the evaluation category of the Nursing Process.

This process was repeated for the remaining four physiological concept categories. Items number seven through 12 pertained to perfusion, 13 through 18 oxygenation, 19 through 24 infection, and 25 through 30 fluid and electrolytes. The first three items in each group were from the framework group and tested cue recognition, nursing action, and evaluation of outcomes using the unfolding case study format in seriation. The remaining three items in each group were from the non-framework group and tested remembering and understanding, analyzing and application, and evaluation. These items functioned independently.

Data Collection

The Dean of the nursing program granted permission to conduct the study at the college using the 2019 second-year cohort. The researcher provided a short PowerPoint presentation that explained the goal of the research study, the process volunteers would undertake to complete the study, and a copy of the consent form (Appendix H). Time was allowed for the researcher to answer questions. The researcher was unknown to the cohort as the researcher was not affiliated with the college where the study was conducted. An invitation was emailed for the researcher by the administrative assistant to the entire cohort requesting volunteers for participation (Appendix
Q). The invitation included the detailed explanation of the reason for the study, emphasizing the upcoming changes to the NCLEX-RN® examination and the value of exposure to the new type of NCLEX-RN® style questions, in hopes of motivating participation.

The most convenient date and time for the students was coordinated with the advanced medical-surgical course faculty member and the Dean. The scheduled date and time, as well as the estimated amount of time to complete the examination, was provided to the students in the invitation. A declaratory statement explained that due to ethical constraints monetary or academic incentives would not be provided but participants would receive a certificate of participation (Appendix R) which they could include in their professional portfolio.

The examination for the study took place in the nursing department’s computer lab using the quiz functionality of the Canvas® Learning Management System in an open access course provided to the researcher by Canvas® (Instructure, 2018). Students were familiar with the computer lab in the nursing department, as well as the Canvas® system because all examinations throughout the nursing program were computerized exams using this computer lab and the Canvas® quiz program. Written permission was obtained from the Dean of the nursing program to conduct the study at this facility (Appendix J). A Canvas® course was provided by Instructure (2018) for the development and administration of the examination.

Volunteers were instructed to simply show up on the day of the examination. The researcher provided the participants with a user ID and password to sign in as an anonymous student to prevent collection of any personal identification data. Once the participants signed in to the Canvas course, the proctor provided the code to access the examination.

Students were allowed to use earplugs if desired to minimize distractions. The college informational technology department was readily available should technical issues occur. A
nursing faculty member who was not directly involved in the students’ program of study proctored the examination. Written instructions were provided to the proctor to ensure replication of the study in the future (Appendix I).

The quiz function of the Canvas® Learning Management System was used to deliver the exam. Students logged in as an anonymous test student using the test student accounts created by the researcher to prevent collection of actual student identification data to maintain anonymity. The first screen contained the consent form (Appendix H) students had to acknowledge in order to access the examination. A written copy of the consent form was also available for each participant. Each exam item was presented one at a time. Each question had to be answered and submitted prior to moving to the next item. Backtracking was not allowed. The final screen contained items for collection of demographical data. The demographical information that was collected included age, race, gender, current GPA range, previously completed college degree or certificate programs, and employment experience in a healthcare environment. All items, including the demographical items, were included in the analysis even if the student decided to submit the exam without completing all the items. Access per participant was set at one-time only to automatically close immediately after the participant submitted the exam to maintain security.

**Data Analysis Methods**

The purpose of this study was to determine the effectiveness of multiple-choice test items in assessing the higher-order cognitive constructs of clinical judgment. The NCSBN® NCJ Task Model was selected to create the multiple-choice test items for this study and were identified as the framework group (Appendix D). The NCSBN® NCJ Task Model provided a framework for test item development that had been proven effective in evaluating the cognitive constructs of
clinical judgment. These constructs were identified by the developers as (a) cue recognition, (b) analysis of cues, (c) prioritize hypotheses, (d) generate solutions, and (e) evaluate outcomes (Dickison et al., 2016).

The first item in the framework group in each physiological concept category, items number one, seven, 13, 19, and 25, assessed the participant's ability to recognize and analyze the assessment data cues presented in the vignette. This item group was labeled the framework assessment group. These items were analyzed using the definitions for the cognitive functions of cue recognition and analysis on layer three of the NCSBN® NCJ Assessment. Participants had to correctly identify and interpret the assessment data provided. Participants had to use critical thinking skills to draw upon nursing knowledge to differentiate between the two pathophysiological findings for the concepts to determine which assessment data would provide the best information upon which to base a nursing plan of action (Dickison et al., 2016; Muntean et al., 2015). Statistical analysis of this group of items using measures of central tendency, frequency distribution and measures of dispersion was performed to determine the individual participant's performance, as well as the overall cohorts' performance. Individual test item performance using levels of difficulty, discrimination index, and point-biserial correlation, as well as sibling test item performance evaluating internal consistency using Cronbach’s alpha was also performed.

The second item in the framework group in each physiological concept categories, items number 2, 8, 14, 20, and 26, assessed the participant's ability to prioritize hypotheses and generate solutions to take the appropriate nursing action. This group of items was labeled the framework implementation group. These items were analyzed using the definitions for the cognitive functions of prioritizing hypotheses and generate solutions to take action which is layer
two of the NCSBN® NCJ Assessment. Participants had to evaluate the assessment findings from the previous item, prioritize client needs, and determine the correct nursing action to implement (Dickison et al., 2016; Muntean et al., 2015). Individual test item performance was conducted using levels of difficulty, discrimination index, and point-biserial correlation, as well as sibling test item performance evaluating internal consistency using Cronbach’s alpha. Statistical analysis of this group of items was also conducted using measures of central tendency, frequency distribution and measures of dispersion to determine the individual participant’s performance, as well as the overall cohorts’ performance. Bivariate correlation was performed to analyze the direction and strength of the relationship between group one and group two. If the results would have had a normal distribution curve, the parametric bivariate correlation test Pearson $r$ would have been used to determine if any significant differences exist between group one and group two. If the distribution curve was skewed, the nonparametric Spearman $\rho$ test would be performed.

The third item in the framework group in each of the physiological concept categories, items number three, nine, 15, 20, and 27, required the participant to evaluate the outcome of the intervention implemented in the previous item to determine therapeutic effectiveness of the nursing action. This group of items was labeled the framework evaluation group. These items were analyzed using the definition of evaluate outcomes to determine if the participant could satisfactorily form the correct clinical judgment which is layer two on the NCSBN® NCJ Assessment Model. Evaluating therapeutic effectiveness of nursing interventions is essential in determining if the client is safe or requires further interventions (Dickison et al., 2016; Muntean et al., 2015). Statistical analysis of this group of items using measures of central tendency, frequency distribution and measures of dispersion was performed to determine the individual
participant’s performance, as well as the overall cohorts’ performance. Individual test item performance using levels of difficulty, discrimination index, and point-biserial correlation, as well as sibling test item performance evaluating internal consistency using Cronbach’s alpha was also performed. Bivariate correlation was performed to analyze the direction and strength of the relationship between group one and group two. If the results would have had a normal distribution curve, the parametric bivariate correlation test Pearson $r$ would be used to determine if any significant differences exist between group one and group two. If the distribution curve was skewed, the nonparametric Spearman $\rho$ test would have been performed.

The comparison group of test items, the non-framework group (Appendix F), for this study were developed using the current NCLEX-RN® test plan recommendations using the NCLEX-RN® client’s needs categories and Bloom’s taxonomy to create short-stemmed test items without seriation (NCSBN, 2015b). These items were very different from the new items that will appear on the Next Generation of the NCLEX-RN® examination in that the information provided in the item stem was limited to one physiological concept only and did not require the student to analyze assessment cues from more than physiological concept and make a clinical decision between two concepts (NCSBN, 2017, Fall). Traditional test item analysis techniques using measures of central tendency, levels of difficulty, discrimination index, as well as test item performance evaluating reliability with point-biserial correlation and internal consistency using Cronbach’s alpha was performed.

The first item in the non-framework group in each of the physiological concept categories, items number four, 10, 16, 21, and 28, required the participant to recognize, recall, and interpret assessment data. This is Bloom’s taxonomy cognitive activities of remembering
and understanding. These items pertained to the Nursing Process step of assessment and were therefore be labeled the non-framework assessment group.

The second non-framework item in each physiological concept group, items number five, 11, 17, 22, and 29, required the participant to differentiate, structure, and implement a nursing action. This is Bloom’s taxonomy levels of analysis and application. These items pertained to the Nursing Process step of implementation and were therefore be labeled the non-framework implementation group. In addition to the traditional item analysis listed above, bivariate correlation was performed to analyze the direction and strength of the relationship between group one and group two. If the results would have had a normal distribution curve, the parametric bivariate correlation test Pearson $r$ would be used to determine if any significant differences exist between group one and group two. If the distribution curve was skewed, the nonparametric Spearman $rho$ test would be performed.

The third non-framework item in each physiological concept group, items number six, 12, 18, 23, and 30, required the participant to check, critique and judge the therapeutic outcome of the nursing interventions. This cognitive function is evaluation of Bloom’s taxonomy and corresponds to the final step in the Nursing Process of evaluation and was therefore labeled the non-framework evaluation group.

In addition to the traditional item analysis listed above, bivariate correlation was performed to analyze the direction and strength of the relationship between this group and group one, as well as this group and group two. If the results would have had a normal distribution curve, the parametric bivariate correlation test Pearson $r$ would be used to determine if any significant differences exist between group one and group two. If the distribution curve was skewed, the nonparametric Spearman $rho$ test would be performed.
The dependent variables were the students answer selections on the exam items. The statistical analysis reports of the results of the examination were generated using the Canvas® reports function. Descriptive and inferential statistics measures as described was used to conduct a quantitative analysis of the data using SPSS data analysis software (IBM, 2018). Charts, tables, and graphs were developed to display the results. The codebook (Appendix S) was developed in Excel and uploaded into the SPSS software for analysis.

The focus of this study was to analyze the student’s answer selections using the NCSBN® NCJ Assessment Model to categorize the results of the framework group using the cognitive operational layers as nominal indicators and then compare these results to the results of the non-framework group. By writing the framework group of test items to conform to the cognitive constructs defined by Dickison et al. (2016), the outcome of clinical judgment was measured using the NCSBN® NCJ Assessment model to analyze the dependent variables. This multi-layered model provided operational layers to the cognitive construct data elements of clinical judgment to allow for nominal measurement of each individual construct (Dickison et al., 2016). Both descriptive and inferential statistics were conducted to analyze the results as follows:

1. Individual test item performance (level of difficulty, discrimination index, and point-biserial correlation, Cronbach’s alpha).
2. Sibling test item performance (Cronbach’s alpha and bivariate correlation).
3. Framework item group (measures of central tendency, frequency distribution measures of dispersion, and comparison of the three framework item groups using bivariate correlation Pearson $r$ if distribution is parametric and Spearman $\rho$ if skewed).
4. Non-framework item group (measures of central tendency, frequency distribution, measures of dispersion, and comparison of the three non-framework item groups using bivariate correlation Pearson \( r \) if distribution is parametric and Spearman \( \rho \) if skewed).

5. Comparison of the results of the framework group to the non-framework group (two-sample \( t \) test if normal distribution; Mann-Whitney \( U \) test if distribution is skewed).

6. Analysis of individual student scores on the framework items per concept (percentage of items answered correctly per physiological concept).

7. Analysis of cohort’s scores on the framework items per concept (percentage of items answered correctly per physiological concept).

Specific detail to the cohorts’ responses to the framework items as compared to the non-framework items was the essential measurement of this study. Statistical analysis was conducted using IBM SPSS 24 software to compare the framework group to the non-framework group. The level of significance that was used to accept or reject the null was an alpha of less than .05.

If the results yielded an alpha less than .05, the null hypothesis would be rejected as the data would indicate there are statistically significant differences in the scores of the multiple-choice test items developed using the NCSBN® NCJ Task Model as a framework, as compared to the multiple-choice test items developed without using a framework. This would provide evidence of a process that could be utilized to develop multiple-choice test items to assess the clinical judgment abilities of prelicensure nursing students.

If the results yielded an alpha greater than .05, this would support the null indicating no statistically significant differences in the scores for the two groups. If the students answered the items in the non-framework correctly but do not answer the framework items correctly, this
would indicate that the students understood basic knowledge of the concepts but could not apply the knowledge to a clinical scenario. Perhaps the students were good test-takers but lacked the ability to form competent clinical judgments.

A second significant continuum measure of this study was the analysis of an individual’s score on the responses to the framework group items. The level of significance that was used to accept or reject the hypotheses is 95% or a $p$-value of < .05. Using the interval rating layer scale on the NCSBN® NCJ Assessment Model, could the nurse educator accurately determine an individual participant’s cognitive operational level of clinical judgment for each of the five physiological concepts for 95% of the participants? The nominal scores of the participants answer selections of the categorical measures were analyzed to determine the student’s interval rating on clinical judgment ability continuum (Appendix E), per item, per concept and on the overall exam (Dickison et al., 2016). The participant had to select the correct answer for all three of the framework items per concept to demonstrate achievement of clinical judgment of the physiological concept to reject the null. This is Layer one of the NCSBN® NCJ Assessment Model.

**Data Management Methods**

The reports generated by the Canvas® Learning Management System were exported to the researcher’s home computer which was password protected and backed up using a password protected external hard drive to prevent loss of data. Both the computer and the hard drive are in the researcher’s private home office. No student identification data was collected as the students signed in using the anonymous test student user ID and password created by the researcher. The IP addresses of the computers used to take the examination were not collected as the settings within the system were set to block.

Electronic copy available at: https://ssrn.com/abstract=3634804
The researcher was the only person who had administrative access to the Canvas® course. The participants’ access was limited by accessibility time and date by the researcher and set to lock after submission of the examination to prevent re-entry. The study was implemented in May 2019. The reports were analyzed immediately thereafter. The raw data will be retained for five years and then deleted from the researcher’s computer and external hard drive.

**Ethical Considerations**

The primary ethical concerns for the participants of this study was to minimize the disruption to the students’ current course work in the nursing program and to maintain the confidentiality of the participants’ exam scores. Information regarding the purpose and process of the study, including how information was handled and disseminated was explained during the recruitment presentation and provided in writing on the informed consent. The consent contained the eight elements identified on the Federal Code for the protection of the participants such as the reason for the study and the potential risks and benefits to the participants (Tappen, 2016, pp. 200-201). Information covering the entire process was provided in the application for approval to the Institutional Review Boards at American Sentinel University and the community college where the study was conducted (Houser, 2015).

The study presented minimal risk to participants. The probability and magnitude of harm or discomfort anticipated during this project were not greater than any ordinarily encountered in daily life. No type of physical risk was induced to the participants by this study. The major psychological risk was anxiety during the examination. Continued assurance of confidentiality and anonymity of the findings was provided to the participants before and after the examination. Re-emphasis of the fact the examination was strictly experimental and will have no bearing on their current course of study was also provided. Students were notified that neither the faculty
member proctoring the examination, nor the Dean of the nursing program would be provided with the individual results of the students’ scores. Participants were reminded prior to the beginning of the examination that their participation was strictly voluntary, and they were free to stop the examination and leave the testing area at any time without fear of any negative consequences.

Internal and External Validity

Rigorous documentation of the specific details of the process of item development was an important component of this study as the quality of the test items pose the highest threat to the internal validity of the findings. The creation of the test items by the researcher may have impacted the feasibility of the test items. To minimize the impact of researcher bias, clinical subject matter experts consisting of three advanced nurse practitioners reviewed each item and provided recommendations for improvement. In addition, three doctoral prepared nurse educators reviewed the items to determine if errors in syntax, grammar, or the use of colloquialisms were present which could distract the participants’ responses during administration of the examination.

Analysis of the results indicating the statistical data and how the results were interpreted was verified with an experienced nurse educator competent in statistical analysis of test item data. The contextual factors identified for this study was clearly delineated in the definition of terms, project design, and methods and procedures section of the written documentation. Specific details of each step of the item development with and without the framework, subject matter expert review and recommendations for improvement, administration of the examination, and analysis of the results was clearly documented in hopes of facilitating critical critique as well as providing an outline for ease of replication.

Electronic copy available at: https://ssrn.com/abstract=3634804
The major external threat to validity was the sincerity of the participation by the subjects. The results of the examination used in this study did not have an impact on the academic standing of the students. As a result, the participants may not have taken the examination seriously and applied their best effort in answering the test items. By passionately explaining the goals of this study to the participants, the researcher hoped to elicit the participants’ sincere efforts in answering the test items.

**Summary**

The consequences of failing to adequately assess the clinical judgment abilities of RN students in undergraduate programs have far-reaching implications. If RN students are unable to pass the NCLEX-RN® examination, they cannot obtain gainful employment which contributes to the current nursing shortage. Novice RNs who cannot form safe clinical judgments can have a negative impact on patient outcomes (Kavanagh & Szweda, 2017; Purling & King, 2012). A possible solution as proposed by this project was to establish an evidence-based method for developing multiple-choice test items to provide an accurate method of assessment of the RN student’s cognitive operational level of clinical judgment. The results of this research add to the body of knowledge in the nursing education literature by providing a validated process to use to consistently develop multiple-choice test items that accurately measure the clinical judgment abilities of undergraduate nursing students.

This quasi-experimental study using a small convenience sample of students from a local community college provided an adequate exploration of the research topic. The results of this research study added to current nursing knowledge by providing insight into a process for developing multiple-choice test items that were structured according to the new NCSBN® item format purpose for the upcoming revisions to the NCLEX-RN® examination (NCSBN, 2017,
Fall). The process delineated in the development of the multiple-choice test items and how to interpret the findings provided an evidence-based method for developing multiple-choice test items and analyzing the results that can be used with large cohorts to measure clinical judgment accurately. Future nurse educators can use the information to aid in the identification of struggling students by analyzing individual student performance. The information could then be used to assist the student in identifying study strategies that would facilitate the development of the higher-order critical thinking skills necessary to apply the theoretical knowledge to a clinical client care situation.
SECTION III: RESULTS AND DISCUSSION OF FINDINGS

Introduction

Registered nurses are responsible for delivering high-quality, competent, evidence-based client care in a variety of healthcare settings (Kavanagh & Szweda, 2017). Of concern is the increasing number of incidents of patient harm or death reported in recent years in the United States (Jessee, 2018; Makary & Daniel, 2016). In response to this crisis, the National Council of State Boards of Nursing has instituted the Next Generation NCLEX-RN® Project to ensure the licensure examination is accurately measuring the candidate’s clinical judgment abilities (NCSBN, 2017, Fall). Nursing faculty in undergraduate nursing programs have a responsibility to assess the RN student’s clinical judgment abilities (National League for Nursing, 2013).

The review of the literature revealed a lack of evidence-based assessment methods for test item construction to assess the cognitive operational levels of clinical judgment for large cohorts of RN students in the didactic setting (Bailey et al., 2012; Bristol et al., 2018; Clifton & Schriner, 2010; Killingsworth et al., 2015; Masters et al., 2001; Tarrant et al., 2006). This DNP capstone project helps fill this knowledge gap by researching the effectiveness of multiple-choice test items in accurately assessing the clinical judgment abilities of the undergraduate RN students. The results will provide evidence upon which to base recommendations for test item construction to nursing faculty in undergraduate RN programs.

In this section of the paper, the methods and procedures used to collect the data will be discussed. Interpretation of the analysis of the findings for the project will be conferred. The characteristics of the sample and the setting in which the study took place will be provided. The data analysis results, as well as the demographic information, will be presented in graphs and tables. Explanation of the implications of the major findings of the project for nursing practice is
expounded upon. Finally, recommendations on how the findings can be used in nursing education and contributions of knowledge to the practice and profession of nursing will be presented.

Summary of Methods and Procedures

A prospective quasi-experimental quantitative comparison design was used to answer the primary research question regarding the difference in the measurement of the Registered Nursing student’s cognitive level of clinical judgment when comparing the results of the student’s answer selections to multiple-choice test items created using the NCSBN® NCJ Model (Dickison et al., 2016) as compared to multiple-choice test items created using the NCSBN® NCLEX-RN Test Plan (NCSBN, 2015b). IRB approval was granted from American Sentinel University as well as the community college at which this study will be conducted. The accessible population selected as the sample ($N = 70$) for this project was the students in their second year of study in an Associate Degree Program in a community college in the Pacific Northwest. The data collection instrument was a 30-item multiple-choice examination consisting of two different groups of item types. Fifteen of the multiple-choice questions (MCQs) developed using the NCSBN Nursing Clinical Judgment Task Model. This group of items was labeled as the framework group. The remaining 15-items were developed using the 2016 NCLEX-RN® Test Plan and were labeled the non-framework group (NCSBN, 2015b). These items were created prior to the publication of the 2019 NCLEX-RN® Test Plan in January 2019 (NCSBN, 2018b), but review of the updates did not reveal any new requirements for test item construction. Therefore, the items created using the 2016 NCLEX-RN® Test Plan remained relevant for use in this study.

Volunteers were recruited to take the examination during the first week of May 2019 by a presentation by the researcher and posting flyers on the bulletin boards in the building where the
students attend class. The examination was administered during the third week of May 2019 as a quiz using the Canvas® Learning Management software in the nursing program’s computer testing lab (Instructure, 2018). A total of 70 volunteers took the examination anonymously. No personal data were collected as the participants were provided with test student user identifications and passwords. The IP addresses were not collected as this setting was set to block. The informed consent was presented as the first page of this electronic exam and required acknowledgement before the exam could be accessed. Demographic information was collected from the participants on the last screen of the electronic exam (Table 1).

Once all participants completed the examination the statistical analysis reports were downloaded from the Canvas® system into an Excel® spreadsheet (Microsoft, 2018) on the researcher’s home computer which is password protected and used only by the researcher. An Excel® codebook was created and uploaded to SPSS® version 24 (2018) for analysis. Reverse coding was not required for any of the variables as none were negatively worded. Recoding of the participants’ answers selections to indicate the incorrect = 0 and correct = 1 responses on each of the MCQs were performed. The Exclude Cases Pairwise setting option in SPSS 24 was used to exclude cases only if the data was missing for the specific analysis. This allowed the case to be included in any of the analyses for which the necessary information was available (Pallant, 2016, p. 58).

Analysis of the demographic data (Figure 1) revealed the majority of the participants were Caucasian (49, 70%) females (59, 84.3%) between the ages of 26 and 30 (22, 31.4%) with a grade point average of 3.5 or above (60, 85.7%), who possessed a degree in a healthcare related field of study (29, 41.4%) and were employed in a healthcare environment (46, 65.7%).
Figure 1. Demographic Data.

The dependent variables of this study were the participants answer selections to the exam items \( n = 30 \). The clinical judgment element descriptors defined by the NCSBN® Nursing Clinical Judgment Assessment Model were used to operationalize the measurement of clinical judgment. The numerical values assigned to each of the cognitive operational layers on the NCJ Assessment Model (Dickison et al., 2016) were used to quantify the measurement of the results of the FW group for analysis. To achieve the clinical judgment layer the participants were required to demonstrate competency in the physiological category by answering all three items in the category correctly.

The same level of measurement was applied to the nonframework group. Participants were required to answer all three items in the physiological category correctly to achieve competency in the category. These items were written at the remember and understand, application and analysis, and evaluation levels of Bloom’s taxonomy as delineated in the NCLEX-RN® test plan (NCSBN, 2015b).
All participants ($N = 70$) answered all items ($n = 30$) on the exam. Participants responded correctly to 29.9% ($n = 448$) of the framework items as compared to 52.3% ($n = 785$) of the nonframework items. Assessment of normality of the framework group revealed the data was asymmetrically distributed ($Mdn = 6.0$, min = 3.0, max = 11.0). The mean was 6.4 ($SD = 1.89$) with a 95% confidence interval [5.94, 6.85]. The 5% trimmed mean of 6.36 and the low standard of error of the mean ($SEM = .23$) indicate extreme scores are not having a strong influence on the mean. The positive skew index (.312, $SE = .287$) indicates the scores are clustered to the left with negative Kurtosis (-.380, $SE = .566$). These findings are supported by the Shapiro-Wilk score ($Sig = .026$) indicating a significant deviation from the normal distribution, as well as the significant Kilmogorov-Smirnov score ($Sig = .007$) suggestive of a violation of the assumption of normality in the distribution of the framework item group scores. This violation is further supported by examination of the histogram which reveals positive skewness and negative kurtosis with clustering of the values to the left. Significant deviation of scores from the expected normal probability line is observed on the Q-Q plot. The detrended Q-Q plot also reveals significant deviation of values from the zero line with clustering of scores below the line on the left. No outliers are present on the box-plot graph (range = 8.0, min = 3.0, max = 11.0, $IQR = 3.0$).

Assessment of normality of the nonframework group revealed the data was asymmetrically distributed ($Mdn = 11.0$, min = 8.0, max = 14.0). The mean was 11.21 ($SD = 1.36$) with a 95% confidence interval [10.89, 11.54]. The 5% trimmed mean of 11.22 and the low standard of error of the mean ($SEM = .163$) indicate extreme scores are not having a strong influence on the mean. The negative skew index (-.082, $SE = .287$) indicates the scores are clustered to the right with negative Kurtosis (-.703, $SE = .566$). These findings are supported by
the Shapiro-Wilk score (Sig = .003) indicating a significant deviation from the normal
distribution, as well as the significant Kilmogorov-Smirnov score (Sig = .001) suggestive of a
violation of the assumption of normality in the distribution of the framework item group scores.
This violation is further supported by examination of the histogram which reveals negative
skewness and negative kurtosis with clustering of the values to the right. Significant deviation of
scores from the expected normal probability line is observed on the Q-Q plot. The detrended Q-
Q plot also reveals significant deviation of values from the zero line with clustering of scores
below the line on the right. No outliers are present on the box-plot graph (range = 6.0, min = 8.0,
max = 14.0, IQR = 2.0).

A chi-square test for independence (with Yates’ Continuity Correction) was used to
explore the relationship between the level of clinical judgment achieved within the two
categorical variables. An independent $t$-test or the Mann Whitney $U$ were thought to be
appropriate for analysis during the initial proposal. However, once the codebook was created
and uploaded to the SPSS system the data did not meet the assumptions required for use of these
tests. The chi-square test for independence was deemed appropriate as this test compares the
frequencies observed between the two variables being measured and provides a crosstabulation
of categorical association measures (Pallant, 2016, p. 218).

The results of the chi-square test for independence (with Yates’ Continuity Correction)
indicated a significant association between the framework group and the nonframework group,
$\chi^2 (1, n = 350) = 76.79, p < .001, \phi = .48$. Complete results are provided in Table 2. Due to the
significant statistical findings ($p < .001$) the null hypothesis was rejected as there is a significant
difference in the measurement of the Registered Nursing student’s cognitive level of clinical
judgment when comparing the results of the student’s answer selections to multiple-choice test

Electronic copy available at: https://ssrn.com/abstract=3634804
items created using the NCSBN® NCJ Model (Dickison et al., 2016) as compared to multiple-choice test items created using the NCSBN® NCLEX-RN Test Plan (NCSBN, 2015b).

The results of the crosstabulation of the framework item group to the nonframework item group revealed 31.8% of participants demonstrated competency in the nonframework group (n = 144), 68.2% failed to demonstrate competency in either group, and zero (0%) participants who failed to demonstrate competency in the nonframework group demonstrated competency in the framework group. However, of the students who did demonstrate competency in the framework group (n = 48), all of these participants also demonstrated competency in the nonframework group (100%).

The competency scores in the nonframework group (n = 144, 41%) were extremely higher than the competency scores in the framework group (n = 48, 14%) indicating the participants understood theoretical knowledge of the physiological concepts but could not apply the theoretical knowledge to the clinical scenario presented in the framework style of multiple-choice questions (Figure 2). Of the percentage of items answered correctly per nursing process category (Figure 3), the highest scores were in the nonframework evaluation group (79%), with the lowest scores occurring the framework assessment group (35%). A comparison of the percentage of items correct per physiological concept per item type (Figure 4), revealed the highest scores were in the nonframework perfusion group (90%) and the lowest scores were in the framework infection group (14%). Examination of the sibling items in the physiological categories in the framework group using Spearman’s rho failed to indicate any significant correlation between the two variables. This lack of correlation is most likely due to the overall poor scores of the cohort on the framework group items which indicates the participants did not understand these concepts.
Individual test item analysis indicated good internal consistency with a Cronbach alpha coefficient of 0.78. Six items demonstrated high limits of difficulty, 19 medium difficulty, and five low difficulty (Figure 5). Point biserial item index correlation (Figure 6) revealed a total of 22 items (73%) demonstrated satisfactory or above discrimination (excellent = 4, good = 9, satisfactory = 9, low = 8, and no = 0). Six items (20%) exhibited ideal functionality as defined by Billings and Halstead (2016, p. 439), with highly discriminating point biserial indexes equal to or greater than three, discrimination indexes between .3 and .8, and having all distractors indicating negative point biserial indexes. These items consisted of:

1. the implementation items in the perfusion category of the framework as well as the nonframework group.
2. the evaluation item in the perfusion framework group.
3. the implementation item in the oxygenation framework group.
4. the evaluation item in the oxygenation nonframework group.
5. and the assessment item in the fluid and electrolyte framework group.

All 70 participants answered all 15 items in the framework group; therefore, the researcher was able to determine all (100%) of the individual participant’s cognitive operation levels of clinical judgment for each of the three categories. The answer selections to each of the assessment, implementation and evaluation items in the five physiological groups for each participant was examined. The participants level of cognitive operation was defined by the item in which the participant scored the lowest. Where participants had two or more categories that were equally low in scores, the lowest layer was considered the participants cognitive operational level. The assessment items represent cue recognition which is Layer 3 on the NCSBN® NCJ Assessment Model (Dickison et al., 2016). The implementation items represent take action
which is Layer 2. The evaluation items are also Layer 2. The total scores indicated that the
participants answered 123 (35%) framework assessment items correctly, 151 (43%) framework
implementation items correctly, and 174 (50%) framework evaluation items correctly. Forty-
eight (69%) of the participants operated at the cognitive operational level of Layer 3, and twenty-
two participants (31%) at Layer 2.

Summary of Sample and Setting Characteristics

The project was conducted at a community college with a convenience sample of the
students in their second year of study. The priori test conducted using the G*Power analysis
software indicated 42 participants were needed to evaluate statistical significance (95% CI, \( p =
.05 \)) between the two groups of independent variables (Faul, Erdfelder, Buchner, & Lang, 2009).
A total of 70 students (\( N = 70 \)) volunteered and participated in the study.

Descriptive statistical analysis was conducted on the demographic information. All 70
participants completed all queries on the demographic survey. The results revealed that the
sample (\( N = 70 \)) consisted of 59 females (84.3%) and 11 males (15.7%). None of the
participants indicated other as gender identification choice.

![Gender of Sample Participants](image)

**Figure 7.** Gender of Sample Participants.
Age categories were queried using the ordinal ranges of 18 to 25, 26 to 30, 31 to 35, 36 to 40, and above 40 years. Descriptive statistical analysis was conducted to evaluate the frequency of the age groups (Figure 8). The results revealed 25.7% \((n = 18)\) participants were from the 18 to 25 category, 31.4% \((n = 22)\) from the 26 to 30 category, 20.0% \((n = 14)\) from the 31 to 35 category, 14.3% \((n = 10)\) from the 36 to 40 category, and 8.6% \((n = 6)\) from the greater than 40 category.

![Age Groups of Sample Participants](image)

**Figure 8.** Age Groups of Sample Participants.

Race of the participants was obtained using the categories of American Indian, Asian, African American, Caucasian and Other. Descriptive analysis of the frequencies of racial categories was conducted (Figure 9). No participants \((0\%, n = 0)\) identified as American Indian or African American. Eleven participants \((15.7\%)\) identified as Asian. Forty-nine participants \((70.0\%)\) identified as Caucasian. Ten participants \((14.3\%)\) identified as Other.
Figure 9. Racial Groups of Sample Participants.

Cumulative grade point averages (GPA) were obtained using ordinal categories of 2.2 to 2.4, 2.5 to 3.0, 3.1 to 3.4, and 3.5 to 4.0. Descriptive analysis was conducted to determine the frequency of each category (Figure 10). Zero participants (0%, $n = 0$) indicated a GPA of 2.2 to 2.4. Three participants (4.3%) indicated a GPA of 2.5 to 3.0. Seven participants (10.0%) cited a GPA of 3.1 to 3.4. Sixty participants (85.7%) reported a GPA of 3.5 to 4.0.

Figure 10. Cumulative GPA of Sample Participants.

Information regarding previous education or training in a healthcare field of study was obtained. Three queries were listed and consisted of No, Yes, non-healthcare field of study, and
Yes, healthcare related field of study. Descriptive analysis was conducted to evaluate the frequencies (Figure 11.) Twenty participants (28.6%) indicated no prior achievements in a healthcare related field of study. Twenty-one participants (30.0%) had achieved a degree or certificate in a non-healthcare related field of study. Twenty-nine participants (41.4%) had obtained a degree or certificate in a healthcare related field of study.

![Previous Healthcare Education of Sample Participants](image)

*Figure 11. Previous Healthcare Education of Sample Participants.*

The final demographical category queried the participants on current or previous employment in a healthcare setting. Three queries were listed and consisted of No, Yes, non-healthcare setting, and Yes, healthcare setting. Descriptive analysis was conducted to evaluate the frequencies. Four participants (5.7%) indicated non-employment status. Twenty participants (28.6%) were currently employed in a non-healthcare setting. Forty-six participants (65.7%) were currently working in a healthcare setting.

**Major Findings**

The data was entered into a codebook for analysis using IBM SPSS® Version 24 (2018). A chi-square test for independence (with Yates’ Continuity Correction) was performed to compare the scores of the participants on the framework test item group and the non-framework item group (Table 2). The results indicated there is a statistically significant association between
the framework group and the nonframework group, $\chi^2(1, n = 350) = 76.79, p < .001, \phi = .48$.

Participants demonstrated a greater level of competency in answering the items created following the NCLEX-RN® Test Plan (NCSBN, 2015b) by selecting the correct answer choices on 52.3% ($n = 785$) of the items, as compared to the items written following the NCSBN® NCJ (Dickison et al., 2016) framework resulting in 29.9% ($n = 448$) of correct answer selections to these items. Since the results yielded an alpha less than .05, the null hypothesis is rejected as the data indicates there is statistically significant difference between the scores of the multiple-choice test items developed using the NCSBN® NCJ Task Model (Dickison et al., 2016) as a framework, as compared to the multiple-choice test items developed using NCLEX-RN® Test Plan (NCSBN, 2015b).

Additional analysis of the individual participant’s scores in each of the five physiological categories was conducted to determine the participant’s cognitive operational level of clinical judgment in each category (Figure 12). Since all participants answered all of the items on the exam, the researcher was able to evaluate 100% of the participants level of clinical judgment. The results indicated that the participants answered all three of the nonframework items correctly within each of the five physiological categories 41% of the time, as compared to only 14% of the time on the framework item. This indicates that the participants understood the theoretical knowledge of the physiological concepts but could not apply the theoretical knowledge to the clinical scenario presented in the multiple-choice questions created using the NCSBN® NCJ framework (Dickison et al., 2016).

Participants demonstrated the highest level of clinical judgment competency in the nonframework perfusion category with 51 participants (72.9%) answering all three items correctly. However, the participants scored extremely low in the framework perfusion category.
with only two participants (2.9%) answering all three items correctly. Similar results are demonstrated in the infection category with 42 participants (60.0%) demonstrating competency in the nonframework group, and only two participants (2.9%) in the framework group. The clinical judgment competency scores were equal in the metabolism group with 21 participants (30%) answering all three items in both groups correctly. Similar scores were demonstrated in the oxygenation group with 25 participants (35.7%) demonstrating competency in the nonframework group, and 22 participants (31.4) in the framework group. The lowest clinical judgment scores occurred in the fluid and electrolyte categories with only four participants (5.7%) achieving competency in the nonframework group as compared to two (2.9%) in the framework group.

![Number of Participants Achieving Clinical Judgment Competency per Physiological Category](image)

*Figure 12. Number of Participants Achieving Clinical Judgment Competency per Physiological Category.*

Results were analyzed to determine the findings using the nursing process theory as a guide to interpret the data (Figure 3). The findings support the relevance of the nursing process theory to nursing practice. Analysis of the results in nursing process category of assessment
revealed 78% \((n = 273)\) of the nonframework items were answered correctly as compared to 35% \((n = 123)\) of the framework items. Assessment is the first step in the nursing process theory and an essential function in the scope of practice of registered nurses. Assessment requires the cognitive activities of recognizing important findings from the subjective and objective components of the physical examination. Assessment represents the cognitive operational function of clinical judgment identified as cue recognition. The results of this analysis reveal participants could recall assessment information when asked a specific question regarding one concept, but could not use the knowledge to recognize pertinent assessment data by differentiating normal from abnormal assessment findings presented for the two concepts in the case scenario.

In the implementation category 68% \((n = 237)\) of the nonframework items were answered correctly and 43% \((n = 151)\) of the framework items. Implementation is the fourth step in the nursing process theory. This step requires the registered nurse to perform the correct nursing interventions to meet the client’s needs. The cognitive activity of application is required to perform the clinical judgment of taking the correction action. The findings reveal participants were more accomplished in determining the correct nursing action with the nonframework items than the framework items.

In the evaluation category participants answered 79% \((n = 275)\) of the nonframework items correctly and 50% \((n = 174)\) of the framework items. Evaluation is the final category of the nursing process. Evaluation is similar to assessment in that the registered nurse must reassess the client and determine the therapeutic effectiveness of the nursing interventions. This step combines the cognitive functions of cue recognition as well as the ability to form sound clinical judgments. The findings reveal participants were able to evaluate the therapeutic effectiveness
when the test item information was presented in the very specific short-stemmed nonframework item, as compared to the unfolding case study style of the framework item group.

The results of this project provide new evidence of a process that can be utilized to develop multiple-choice test items to assess the clinical judgment abilities of prelicensure nursing students. The review of the literature indicated a lack of evidence-based item writing methods to guide the development of items that can accurately assess the higher-cognitive constructs required for sound clinical judgments in prelicensure registered nursing educational programs (American Nurses Association [ANA], 2015a; Bristol & Brett, 2015; Bristol et al., 2018; Clifton & Schriner, 2010; Jesse, 2018; Killingsworth, 2013; Masters et al., 2001; Muntean, 2012; NCSBN, 2018a, January; Stanton, 1983; Sutherland et al., 2012; Tarrant et al., 2006; Tarrant & Ware, 2008). The findings of this study support the use of the NCSBN® NCJ Task model (Dickison et al., 2016) as a framework to use to develop multiple-choice questions that can measure the clinical judgment abilities of the nursing students.

The case study information presented in the item stems of the framework item group represented realistic assessment information for client’s experiencing hypoglycemia, hemorrhagic shock, oxygen deprivation, septic shock, and fluid overload. The low percentage of achievement (35%) in recognizing the early signs and symptoms indicating impending client demise in these items is of concern. The findings on the lack of ability to recognize early assessment information support the information revealed in the review of the literature from studies examining recognition of assessment findings and the impact of the RNs clinical judgment ability on providing competent patient care (Di Vito-Thomas, 2000; Jesse, 2018; Kavanagh & Szweda, 2017; NCSBN, 2015a; Purling & King, 2012). The results of this study provide new evidence on a method to create multiple-choice questions that specifically assess the
components of the nursing process. This information adds to the body of knowledge on test item development and how the analysis of the test data can be used to evaluate the individual student’s level of clinical judgment in each step of the nursing process.

**Implications for Nursing Practice**

Limited research is available that provides evidence of best practices for writing multiple-choice test items. No research findings are available that specify a framework that can be utilized to create multiple-choice questions to assess the clinical judgment abilities of registered nursing students. Clinical judgment is a critical component of the scope of practice of registered nurses essential in providing competent patient care. The purpose of this project was to evaluate the effectiveness of multiple-choice test items developed using the NCSBN® NCJ Task Model (Dickison et al., 2016) in accurately assessing the clinical judgment abilities of prelicensure registered nursing students as compared to the results from multiple-choice test items developed using the current NCSBN® NCLEX-RN test plan (NCSBN, 2015b).

The results of the data analysis conducted for this project indicated statistical significance in the evaluation of the two different types of test items. The results of this project provide evidence that the NCSBN® NCJ Task Model (Dickison et al., 2016) can be used as a framework to create multiple-choice test items that can be analyzed to accurately determine the cognitive operational levels of clinical judgment as defined by the NCSBN® on the NCJ Assessment Model (Dickison et al., 2016). The findings of the project support the claim of Dickison et al. (2016) that multiple-choice questions can be developed to assess the higher order cognitive operations of critical thinking necessary to form clinical judgments. The results of this study dispute the contrasting reports by Masters et al. (2001) that multiple-choice questions can only measure the lower levels of cognitive operations defined by Bloom’s taxonomy. Due to the
upcoming changes proposed by the NCSBN® with the Next Gen NCLEX project the findings of this study provides relevant information for nurse leaders and nurse educators to examine test item development and item analysis policies.

**Recommendations**

The significance of this capstone project was to explore the effectiveness of multiple-choice test items in evaluating clinical judgment. The results indicated a significant difference in the cognitive operational levels of the prelicensure registered nursing students when answering the nonframework test items as compared to the framework test items. Evidence-based methods are needed that provide a framework for the development of high-quality multiple-choice test items that provide data that can be analyzed to determine the students level of cognitive function. Nurse educators can benefit from the findings of this study as a basis for incorporating the use of the NCSBN® NCJ Task model (Dickison et al., 2016) as a framework to guide the development of test items. Specific attention to the criteria of incorporating more than one clinical concept into the item stem in order to assess the student’s ability to recognize, analyze, prioritize and formulate competent clinical judgments is a major recommendation of this study. The results of the statistical significant difference in the items using a vignette to present realistic client assessment data pertaining to two or more physiological systems supports the recommendation of Case, and Swanson (2003), Coderre, Harasym, Mandin, and Fick (2004), Johnson (1999) and Su, Osisek, Montgomery, and Pellar (2009).

The findings indicating the lack of cue recognition of abnormal assessment data demonstrates how the framework items can be analyzed to identify the specific cognitive operational level of clinical judgment that is lacking. This information can be used to advise the individual student in the development of a study plan that focuses on the specific criteria
identified from the exam. This information can also be used by the nurse educator to identify deficiencies in the cohort in order to develop teaching strategies that will target the areas needing improvement in hopes of improving the groups’ cognitive functional abilities in the specific area of concern.

Administrators in undergraduate nursing educational programs may find the results of this study beneficial in guiding the development of test item writing and test analysis policies. The use of the NCSBN® Nursing Clinical Judgment model may prove beneficial in limiting the impact of a decline in first time pass rates due to the upcoming changes on the NCLEX-RN® examination as indicated in the Next Generation project. Incorporating the processes identified in this project that follow the guidelines for test item development from the NCSBN® may be advantageous in improving the assessment of registered nursing students.

**Recommendations for Future Research**

This capstone project helps fill in the gap of the lack of evidence-based exploration of multiple-choice test items and the assessment of clinical judgment. Nurse educators need proven methods to draw upon to ensure accurate assessment of the clinical judgment capabilities of the registered nursing students in the didactic courses. This project can serve as a beginning framework for future studies to continue to evaluate best practices for developing multiple-choice test items for use in undergraduate nursing programs that specifically focus on the assessment of clinical judgment. More studies are needed with larger, more diverse populations in a variety of undergraduate nursing educational programs. One recommendation would be to create a massive online open access course that can be distributed to nurse educators in undergraduate nursing programs throughout the country. This would provide information from
all types of undergraduate programs from numerous locations with a variety of demographics characteristics.

**Discussion**

The results of this capstone study are encouraging. The framework provided by the NCSBN® was beneficial in the development and analysis of the test items. The differences in the measurement of clinical judgment between the two item groups were very apparent. The large effect size \( d = .5 \) indicates practical significance between the two types of test items.

A strength of the study was the attention in documenting the process used to develop the different types of test items. The findings provide evidential support for the use of the framework items in measuring the cognitive operations necessary for clinical judgment. By documenting the process taken to create the framework items the process can be replicated in future studies. This was important because a key aspect of this project was to explain a process that nurse educators can consistently use to guide the development of test items and to explore the effectiveness of these items in measuring clinical judgment as defined by the NCSBN®. The information revealed from the review of the literature indicated the study by Dickison et al. (2016) provided a working model, the NCSBN® Nursing Clinical Judgment Task Model, that specified a process to create multiple-choice test items to measure the higher-order cognitive operational constructs of clinical judgment. The foundational concepts used to guide the development of the test items can be attributed to the constructs of Orlando’s Nursing Process theory (Schmieding, 2002) and the cognitive operational concepts of clinical judgment identified by Dickison et al. (2016). Orlando’s constructs regarding the autonomous deliberative nursing actions based on the perceptions of the observations during the nurse-patient encounter are supported by the findings of this study as well as the original work of Muntean’s study (2012)
and subsequent supportive study by Dickison et al. (2016) of the operational definitions of the constructs of clinical judgment. The realistic client assessment data provided in the item stems covering more than one physiological concept supports the continued relevance of the Nursing Process Theory which continues to be utilized as a theoretical framework both in nursing education and nursing practice. The findings from this study regarding the difference in measurement between the framework assessment, implementation, and evaluation item groups as compared to the non-framework item groups covering the same concepts supports the findings of the continued relevance of these conceptual foundations of the NPT as indicated in the studies by Shabel (1989), Potter and Bockenhauer (2000), and McEwen and Brown (2002).

Another strength of this study is the large amount of data obtained anonymously from the convenience sample which consisted of students unknown to the researcher. Bias that could occur from being influenced by familiarity of the participants was not an issue. In addition, the large data sample ensured a sample size large enough to minimize the potential of a type I or type II error.

A limitation of this study is the use of a single associate degree nursing program of study at one community college. More research needs to be conducted to incorporate findings from all types of undergraduate nursing programs including baccalaureate programs. The majority of the sample were of Caucasian ethnicity and female gender. Expansion of the study to reach a more diverse population would be beneficial in assessing a sample more representative of the registered nursing student population.

Another limitation of this study is the creation of the items by one person. Although the test items were reviewed by a subject matter expert panel, exploration of test item development by faculty teams should be explored. This recommendation would support a change in nursing
educational practice as indicated by the lack of importance placed on peer review of faculty developed test items found from the results of the study conducted by Killingsworth et al. (2015). However, due to the wide-spread use of multiple-choice questions throughout undergraduate nursing programs as indicated by the studies of Clifton and Schriner (2010) and Tarrant, Knierim, Hayes, and Ware (2006), an emphasis on faculty development and peer review oversight is recommended. Due to the ethical responsibilities of nurse educators to ensure fairness in testing and the implications of poorly constructed test items found by White and Heitzler (2018), the findings of this project support the recommendations of Bristol, Nelson, Sherrill, and Wangerin (2018) for nursing programs to create policies for test item development and analysis that incorporates a team approach.

**Conclusion and Contributions to the Profession of Nursing**

Jesse (2018) identified early recognition of problematic cues and effective implementation of evidence-based nursing interventions as key factors in preventing patient injury and death. The underlying concepts of the Nursing Process Theory (Schmieding, 2002) incorporated into the RN Scope of Practice (ANA, 2015a) of assessment, implementation, and evaluation continue to have a significant impact on the quality of client care. The NCSBN® Practice Analysis Reports have indicated the majority of the tasks performed by registered nurses require the ability to form competent clinical judgments (NCSBN, 2018, January). Nurse educators serving as faculty in undergraduate registered nursing programs have the responsibility of assessing the clinical judgment abilities of the students in the program (NLN, 2013). Bristol, Nelson, Sherrill, and Wangerin (2018) discuss the lack of evidence defining best practices for test item development and analysis and call for exploration of evidence-based strategies to construct valid and reliable multiple-choice test items as this is relevant to current practices in
nursing education. This capstone project contributes evidence to the body of knowledge regarding the effectiveness of multiple-choice test items. More exploration is needed in this area to provide evidence-based methods for test item construction nurse educators can use to accurately assess the clinical judgment of registered nursing students. Improvement in the clinical judgment abilities of novice registered nurses relies upon improving the quality of assessment of the undergraduate registered nursing students.
References


Bristol, T., & Brett, A. L. (2015). Test item writing: 3Cs for successful tests. Teaching and Learning in Nursing, 10, 100-103. doi:10.1016/j.teln.2015.01.004


Gillespie, M., & Peterson, B. L. (2009). Helping novice nurses make effective clinical decisions: The situated clinical decision-making framework. *Nursing Education Perspectives, 30*(3),


Hooper, J.I. (2017, November). An evidence-based cognitive approach for teaching nursing clinical judgment skills to entry-level practitioners. Presented at Sigma Theta Tau
International, the Honor Society of Nursing, 44th Biennial Convention, Indianapolis, IN.

Presentation slides retrieved from

http://www.nursinglibrary.org/vhl/handle/10755/623121


Johnson, G. J. (2013). The nursing process and allegations of malpractice against perinatal
registered nurses in Florida (Doctoral dissertation). Available from ProQuest Dissertations & Theses database. (UMI 3577284)


Retrieved from https://search.proquest.com/docview/203965313?accountid=169658


Electronic copy available at: https://ssrn.com/abstract=3634804


Electronic copy available at: https://ssrn.com/abstract=3634804


http://www.raosoft.com/samplesize.html


Scan Tron, Inc. (2019). *Optical grading solutions*. Eagan, MN:

https://www.scantron.com/company/


doi:10.1097/NNE.0b013e3181b2b546


## Tables

### Table 1

**Demographic data**

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency (N = 70)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-25</td>
<td>18</td>
<td>25.7</td>
</tr>
<tr>
<td>26-30</td>
<td>22</td>
<td>31.4</td>
</tr>
<tr>
<td>31-35</td>
<td>14</td>
<td>20.0</td>
</tr>
<tr>
<td>36-40</td>
<td>10</td>
<td>14.3</td>
</tr>
<tr>
<td>&gt;40</td>
<td>6</td>
<td>8.6</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>11</td>
<td>15.7</td>
</tr>
<tr>
<td>Female</td>
<td>59</td>
<td>84.3</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Asian</td>
<td>11</td>
<td>15.7</td>
</tr>
<tr>
<td>African American</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Caucasian</td>
<td>49</td>
<td>70.0</td>
</tr>
<tr>
<td>Other</td>
<td>10</td>
<td>14.3</td>
</tr>
<tr>
<td><strong>GPA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2-2.4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2.5-3.0</td>
<td>3</td>
<td>4.3</td>
</tr>
<tr>
<td>3.1-3.4</td>
<td>7</td>
<td>10.0</td>
</tr>
<tr>
<td>3.5-4.0</td>
<td>60</td>
<td>85.7</td>
</tr>
<tr>
<td><strong>Degree</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>20</td>
<td>28.6</td>
</tr>
<tr>
<td>Non-healthcare</td>
<td>21</td>
<td>30.0</td>
</tr>
<tr>
<td>Healthcare</td>
<td>29</td>
<td>41.4</td>
</tr>
<tr>
<td><strong>Employment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>4</td>
<td>5.7</td>
</tr>
<tr>
<td>Non-healthcare</td>
<td>20</td>
<td>28.6</td>
</tr>
<tr>
<td>Healthcare</td>
<td>46</td>
<td>65.7</td>
</tr>
</tbody>
</table>
Table 2

SPSS Output Chi Square for Independence Test

Case Processing Summary

<table>
<thead>
<tr>
<th></th>
<th>Cases</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Valid</td>
<td>Missing</td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>Percent</td>
<td>N</td>
<td>Percent</td>
<td>N</td>
<td>Percent</td>
</tr>
<tr>
<td>FW_CJ *</td>
<td>350</td>
<td>100.0%</td>
<td>0</td>
<td>0.0%</td>
<td>350</td>
<td>100.0%</td>
</tr>
<tr>
<td>NFW_CJ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FW_CJ * NFW_CJ Crosstabulation

<table>
<thead>
<tr>
<th></th>
<th>NFW_CJ</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 Not Competent</td>
<td>1 Competent</td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FW_CJ</td>
<td>Count</td>
<td>% within FW_CJ</td>
<td>Adjusted Residual</td>
<td>0 Not Competent</td>
<td>Count</td>
<td>% within FW_CJ</td>
</tr>
<tr>
<td></td>
<td>206</td>
<td>68.2%</td>
<td>8.9</td>
<td>96</td>
<td>31.8%</td>
<td>-8.9</td>
</tr>
<tr>
<td></td>
<td>1 Competent</td>
<td>0.0%</td>
<td>-8.9</td>
<td>48</td>
<td>100.0%</td>
<td>8.9</td>
</tr>
<tr>
<td>Total</td>
<td>Count</td>
<td>206</td>
<td>144</td>
<td>350</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>% within FW_CJ</td>
<td>58.9%</td>
<td>41.1%</td>
<td>100.0%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Chi-Square Tests

<table>
<thead>
<tr>
<th>Value</th>
<th>df</th>
<th>Asymptotic Significance (2-sided)</th>
<th>Exact Sig. (2-sided)</th>
<th>Exact Sig. (1-sided)</th>
</tr>
</thead>
</table>

Electronic copy available at: https://ssrn.com/abstract=3634804
<table>
<thead>
<tr>
<th>Test</th>
<th>Value</th>
<th>Approximate Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>79.581</td>
<td>.000</td>
</tr>
<tr>
<td>Continuity Correction</td>
<td>76.789</td>
<td>.000</td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>96.504</td>
<td>.000</td>
</tr>
<tr>
<td>Fisher's Exact Test</td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td>Linear-by-Linear Association</td>
<td>79.353</td>
<td>.000</td>
</tr>
</tbody>
</table>

N of Valid Cases: 350

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 19.75.
b. Computed only for a 2x2 table

**Symmetric Measures**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
<th>Approximate Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal by Nominal Phi</td>
<td>.477</td>
<td>.000</td>
</tr>
<tr>
<td>Cramer's V</td>
<td>.477</td>
<td>.000</td>
</tr>
</tbody>
</table>

N of Valid Cases: 350
Figures

Figure 2. Total number of competencies achieved per category.
Figure 3. Percentage of items correct for nursing process categories of assessment (A), implantation (I), and evaluation (E) in the framework (FW) and nonframework (NFW) categories.
Figure 4. Comparison of percentage of items correct per physiological concept per item type of metabolism (M), perfusion (P), oxygenation (O), infection (I), and fluid and electrolytes (FE) per framework (FW) and nonframework (NFW) groups.
Figure 5. Item level of difficulty index. Less than three indicates items of high difficulty. Items of medium difficulty range between 0.3 to 0.8. Items considered easy, of low difficulty are greater than 0.8 (Billings and Halstead, 2016, pp. 438-439).
Figure 6. Point biserial discrimination index of test items. Items with excellent discrimination have a point biserial index (PBI) of four or greater. Items exhibiting good discrimination have a PBI between 0.3 and 0.39. Satisfactory discrimination is indicated by a PBI of 0.15 to 0.29. Items with low discrimination have a PBI of <0.15. Zero test items had a PBI of zero indicating no discrimination.
Appendix A

NCSBN® Nursing Clinical Judgment Assessment Model

Appendix B

NCSBN® Nursing Clinical Judgment Task Model

<table>
<thead>
<tr>
<th>Cognitive Operation</th>
<th>Factor Conditioning</th>
<th>Expected Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognize Cues</td>
<td>Environmental Cues: Set location to emergency room&lt;br&gt; Show the presence of parent&lt;br&gt; Patient Observation Cues: Show age to 8-10&lt;br&gt; Show dehydration symptoms (e.g., dry mucous membranes appear, cool extremities, cap refill 3-4 seconds)&lt;br&gt; Show/Imply lethargy&lt;br&gt; Medical Record Cues: Show dehydration symptoms (e.g., a lower-grade temperature, diarrhea, a poor appetite)&lt;br&gt; Show/Imply history of diabetes&lt;br&gt; Show/Imply vital signs&lt;br&gt; Time Pressure Cue: Set time pressure to varying with onset of symptoms and current lethargy</td>
<td>Recognize abnormal vital signs&lt;br&gt; Recognize symptoms of dehydration&lt;br&gt; Identify the history of diabetes&lt;br&gt; Hypothesize dehydration&lt;br&gt; Hypothesize diabetes</td>
</tr>
<tr>
<td>Analyze Cues</td>
<td>Require knowledge of dehydration symptoms&lt;br&gt; Require knowledge of diabetes symptoms</td>
<td></td>
</tr>
<tr>
<td>Prioritize Hypotheses</td>
<td>Give vital sign monitors as resources&lt;br&gt; Set time pressure to vary with vital signs</td>
<td>Prioritize dehydration&lt;br&gt; Address dehydration&lt;br&gt; Avoid glucose</td>
</tr>
<tr>
<td>Generate Solutions</td>
<td>Require knowledge of dehydration treatment and intervention&lt;br&gt; Require knowledge of diabetes treatment and intervention</td>
<td></td>
</tr>
<tr>
<td>Evaluate Outcomes</td>
<td>Experience:&lt;br&gt; Require experience of administering isotonic fluid&lt;br&gt; Patient Observation Cue:&lt;br&gt; Show patient awaking and talking&lt;br&gt; Imply &lt;Set vital signs to varying with action&gt;</td>
<td>Check vital signs&lt;br&gt; Check lethargy</td>
</tr>
</tbody>
</table>
Appendix C

Relationship of Orlando’s Nursing Process, Bloom’s Cognitive Activities, and NCSBN® NCJ Cognitive Operations and Assessment Layers

(Angerson et al., 2001; Dickison et al. 2016; Schmieding, 2002)
Appendix D

Multiple-Choice Test Items – Framework Group

<table>
<thead>
<tr>
<th>Test Item Number</th>
<th>Codebook Variable Name</th>
<th>Item</th>
</tr>
</thead>
</table>
| 1                | M_FW_A                 | Item #1 –  
                  | Primary concept - metabolism  
                  | Secondary concept – perfusion  
                  | A client’s daughter calls the nurse’s station requesting help for her mother. She states her mother suddenly doesn’t know where she is and her speech is difficult to understand. The client is 3 hours post-operative left total hip replacement. She was transferred from the recovery room to the medical/surgical/telemetry unit 1-hour ago from the post-anesthesia care unit. Upon arrival to the unit she was alert and oriented to person, place, time and situation. Vital signs have been stable: respirations 14, apical pulse 76 irregular rhythm, blood pressure 132/88.  
                  | Previous medical history: Type 2 diabetes mellitus, hypertension, right cerebral vascular accident 5-years ago, and atrial fibrillation.  
                  | Assessment at this time reveals: Vital signs: respirations 12, apical pulse 92 irregular rhythm, blood pressure 134/86. Speech is slurred. She is holding her head and nods yes when asked if she has a headache. She is oriented to name only, appears restless, trying to get out of bed. Moves all extremities well except left leg. Left hip dressing remains clean, dry and intact. Pedal pulses 2+ bilaterally. Hand grasp strong and equal bilaterally. Peripheral intravenous access device 22 gauge in right forearm infusing Ringers Lactate at 125 mL/hr.  
                  | Which of the following assessments should the nurse conduct first?  
                  | a) NIH stroke scale  
                  | b) Analyze telemetry rhythm  
                  | c) **Check the capillary blood glucose level**  
                  | d) Circulation, motion and sensation of left lower extremity  
                  | Rationale – Objective assessment findings indicative of hypoglycemic episode include change in level of consciousness and tachycardia (Hendricks, 2018, p. 1309). Subjective assessment findings include a history of DM and status post-surgical procedure. Finding differentiating hypoglycemia from perfusion problem include strong
<table>
<thead>
<tr>
<th>Test Item Number</th>
<th>Codebook Variable Name</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>equal hand grasp, tachycardia without hypotension, and dry surgical dressing. Irregular heart rhythm is expected with atrial fibrillation.</td>
</tr>
</tbody>
</table>
| 2                | M_FW_I                 | Assessment findings:  
NIH stroke scale score = 5  
Telemetry rhythm = irregular rhythm, rate 108  
Capillary blood glucose level = 38 mg/dL  
Lower extremities warm to touch, capillary refill less than 3 seconds  
What is the priority nursing intervention?  
   a) **Administer 50 mL of D50W intravenous push**  
   b) Administer oxygen 2 liters per nasal cannula  
   c) Administer adenosine 1 mg IV bolus over 1 minute  
   d) Administer heparin 10,000 units subcutaneously  
Rationale – D50W will increase serum glucose levels. Administration of glucose by oral route should be avoided in client’s who may experience difficulty swallowing. Slurred speech key assessment cue (Hendricks, 2018, p. 1311). NIH stroke scale 5 out of 40 does not indicate cerebral injury so heparin therapy would not be indicated (Willis, 2018, p. 931-932). Adenosine is given for paroxysmal supraventricular tachycardias and is not used to treat atrial fibrillation (Dechant & Heimgartner, 2018, p. 677-679). Administration of oxygen would not harm the client, but also would not address CBG of 38 mg/dL. |
| 3                | M_FW_E                 | Which of the following is an early indicator that the intervention the nurse performed is therapeutically effective?  
   a) Pulse oximeter reading 96%  
   b) Heart rate 88 regular rhythm  
   c) **Alert and oriented to person, place and time**  
   d) Pedal pulses to left lower extremity 2+  
Rationale – The central nervous system is sensitive to change in serum glucose levels because the brain cannot produce or store glucose. The neurons must have a steady supply to prevent dysfunction and cellular death. An early symptom of hypoglycemia or hyperglycemia is change in neurological status such as irritability or confusion. When BG levels return to normal limits CNS function improves (Hendricks, 2018, pp. 1281, 1282, 1311). |
<table>
<thead>
<tr>
<th>Test Item Number</th>
<th>Codebook Variable Name</th>
<th>Item</th>
</tr>
</thead>
</table>
| 7                | P_FW_A                 | Item #2 –  
Primary concept - perfusion  
Secondary concept – metabolism  

Upon admission to the medical/surgical unit status post laparoscopic cholecystectomy the client’s assessment finding are as follows: client is sleepy, but easily responds to verbal stimuli, oriented to person, place, time and situation. Respirations 12, blood pressure 128/72, apical pulse 82, temp 37.2 C (98.9 F), oxygen saturation level 98% on room air. Steri-strip dressings clean, dry and intact to 5 laparoscopic abdominal incisions. Abdomen is soft, nondistended and tender to touch. Bowel sounds are absent in all 4 quadrants. Nasal gastric tube is patent and attached to low intermittent suction. Client is due to void. Pain at this time 1/10. Peripheral intravenous access device 20 gauge in right hand infusing D5 0.45% sodium chloride at 125 mL/hr.  

Four hours after admission the client reports abdominal pain 5/10. She states she is nauseated and is shaking. Assessment reveals: Respirations 22, blood pressure 110/52, apical pulse 108, temp 36.5 C (97.7 F), skin cool, diaphoretic.  

Which of the following assessments should the nurse conduct first?  

- a) Capillary blood glucose level  
- b) Mean arterial blood pressure  
- c) Abdominal assessment  
- d) Assess nasogastric tube for patency  

Rationale – Objective assessment cues indicative of hypovolemic shock include tachypnea, tachycardia, hypotension, diaphoresis, decrease in temperature with cool skin (Heimgartner, 2018a, pp.756-758). These are relevant changes from the client’s previous readings indicating a change in the client’s baseline physiological status. Subjective assessment includes status post-surgical procedure. Monitoring of the mean arterial blood pressure is key component in the treatment of hypovolemic shock. A decrease in the MAP of 10 mm Hg from baseline without known cause (such as administering an antihypertensive medication) warrants investigation to determine the cause. Nausea, shaking, and tachycardia are signs of hypoglycemia but should be ruled out due to the hypotension. Abdominal assessment would be needed but is not the top priority as physiological indicators need to be addressed first.
<table>
<thead>
<tr>
<th>Test Item Number</th>
<th>Codebook Variable Name</th>
<th>Item</th>
</tr>
</thead>
</table>
| 8                | P_FW_I                 | Assessment findings:  
  Capillary blood glucose level = 112 mg/dL  
  Mean arterial blood pressure = 71 mmHg down from 91 mmHg on admission  
  Abdominal = slightly distended, painful to palpation, incisions clean, dry, intact  
  Nasogastric tube patent to low intermittent suction with 80 mL dark green secretions  
  What is the priority nursing intervention?  
  a) Administer D50W intravenous bolus  
  b) Administer 500 mL 0.9% sodium chloride intravenous bolus  
  c) Administer meperidine 50 mg IV  
  d) Administer phenergan 12.5 mg IV and offer ice chips  
  Rationale – A 20-point decrease in MAP indicates a significant change in vascular volume which requires replacement to sustain perfusion of vital organs (Heimgartner, 2018a, pp.756-759). CBG is within normal range so D50W is not needed. Administration of narcotic and/or antiemetic medication for pain is contraindicated at this time due to hypotension. Client should remain NPO as surgical intervention may be needed. |
| 9                | P_FW_E                 | Which of the following is the most important finding that the client’s condition is **deteriorating** and the intervention the nurse performed is **not** having the expected therapeutic effect?  
  a) Pain level reported 4/10  
  b) Oxygen saturation level 92%  
  c) Mean arterial pressure 68 mmHg  
  d) Client vomits 250 mL greenish red emesis  
  Rationale – Further decline in the MAP reading indicates the client’s vascular volume is continuing to decline. Additional volume replacement with blood products and drug therapy with vasoconstrictor agents should be considered (Heimgartner, 2018a, p. 759). |
| 13               | O_FW_A                 | Item #3 –  
 Primary concept - oxygenation  
 Secondary concept – fluid and electrolytes |
<table>
<thead>
<tr>
<th>Test Item Number</th>
<th>Codebook Variable Name</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A 72-year-old female client is admitted to the telemetry unit in respiratory distress. Client is alert and oriented person, place, time and situation. Respiration rate 22, leaning forward, labored. Oxygen saturation level 84% on room air. Radial pulse 112. Blood pressure 158/96, denies chest pain. Very poor appetite for the past 7 days. She states she is “just too tired to eat.” Gait is steady, but unable to ambulate more than a few feet due to severe dyspnea on exertion. Medical history: Type 2 diabetes mellitus, chronic obstructive pulmonary disease, hypertension, atrial fibrillation</td>
</tr>
</tbody>
</table>

New medications to begin on admission:
- Furosemide 80 mg IVP STAT one time only
- Furosemide 40 mg IVP every 12 hours
- Heparin loading dose 10,000 units subcutaneously one time only
- Heparin maintenance dose 5,000 units every 12 hours X 3 days then discontinue
- Digoxin digitalizing dose 1 mg IV one time only
- Digoxin maintenance dose 0.125 mg po every day
- Warfarin 5 mg po every day for 4 days, then 2 mg po every day
- Insulin to sliding scale

Previous medications to be continued:
- Lisinopril 40 mg po daily
- Glyburide 5 mg po every day
- Tiotropium 18 mcg by inhaler daily

Which of the following assessments should the nurse conduct first?

a) Obtain weight using bedside scale
b) Arterial blood gases
c) Telemetry reading
d) Auscultate lungs

Rationale – Objective assessment cues indicative of respiratory distress include posturing, tachypnea, hypoxemia, tachycardia, and dyspnea (Workman, 2018, pp. 574-576). Subjective assessment findings include history of COPD, respiratory medications. All assessments should be performed but auscultation of lungs is needed to establish baseline assessment findings, as well as to determine if client is experiencing pulmonary edema from fluid overload or bronchial
<table>
<thead>
<tr>
<th>Test Item Number</th>
<th>Codebook Variable Name</th>
<th>Item</th>
</tr>
</thead>
</table>
| 14               | O_FW_I                 | Assessment findings:  
                      Weight 84 kg  
                      Arterial blood gas: pH = 7.32, pCO² = 52 mmHg,  
                      HCO³ = 36 mEq/L, pO² = 79 mmHg  
                      Telemetry – rate 114, irregular, atrial fibrillation with occasional premature ventricular contractions  
                      Auscultation of lungs - course crackles throughout all lobes anteriorly and posteriorly  
                      Cough productive copious amount of frothy white sputum |
|                  |                        | What is the priority nursing intervention?  
                      a) **Furosemide 80 mg IVP**  
                      b) Tiotropium 18 mcg by inhaler  
                      c) Digoxin digitalizing dose 1 mg IV  
                      d) Heparin loading dose 10,000 units subcutaneously |
| 15               | O_FW_E                 | The client receives all medications ordered over the next 24-hours. Which of the following assessment findings is the best indication of the therapeutic effectiveness of the treatment plan?  
                      a) Heart rate 88  
                      b) Blood pressure 130/72  
                      c) Respiration rate 16  
                      d) Oxygen saturation level 96% |
| 19               | I_FW_A                 | Item #4 – Primary concept – infection (sepsis)  
                      Secondary concept – oxygenation |
A 28-year old client was transferred from the intensive care unit to the medical-surgical unit four hours ago. Admitting diagnoses: multiple injuries sustained in motor vehicle accident 5 days ago, ORIF left humerus, ORIF left femur, fractured left 7, 8 & 9 ribs, pneumothorax left lung (resolved – chest tubes removed 1 hour before transfer to medical-surgical unit).

Assessment at this time reveals:
Vital signs: temp 38.8 C (101.8 F), radial pulse 98, resp 24, oxygen saturation level 90% on room air, blood pressure 108/52. Oriented to person and place. Irritable and restless. Pain 2/10. Skin is warm to touch. Surgical incisions left arm and left leg clean, no drainage, staples intact, no redness or edema. Chest tube insertion site dressings clean, dry and intact. Central line in the upper left chest wall with D5 0.45%NS infusing at 125 mL/hr. Foley urinary catheter patent with 100 mL clear yellow urine.

Which of the following assessments should the nurse conduct first?

a) Arterial blood gases
b) Complete blood count
c) Serum glucose level
d) Serum lactate level

Rationale – Objective assessment cues indicative of infectious process with early symptoms of sepsis include elevated temperature, tachycardia, tachypnea, hypotension, hypoxemia, and low urinary output 25 mL/hr for past 4 hours despite receiving 125 mL/hr IV fluid (Heimgartner, 2018a, pp. 761-763). Subjective assessment findings include recent surgical procedures for traumatic injuries and 3 different types of line – chest tubes, central line and urinary catheter. Findings differentiating from simple oxygenation problem include decline in urinary output, elevated temperature and hypotension. Answer choices a, b, and c will require specimens to be sent to laboratory which is time consuming. Serum lactate test is a point-of-care test that can be completed quickly at bedside. Early recognition of sepsis and intervention is essential for survival (p. 765).

20 I_FW_I 

Assessment findings:
Arterial blood gas: pH = 7.32, pCO₂ = 52 mmHg, HCO₃⁻ = 20 mEq/L, pO₂ = 69 mmHg
Complete blood count: WBC 18,000, Bands 12%, RBC 4.8, Hemoglobin 14.0, Hematocrit 38%, Platelet count 98,000
Serum glucose level = 340 mg/dL
Serum lactate level = 4 mmol/L (36 mg/dL)
<table>
<thead>
<tr>
<th>Test Item Number</th>
<th>Codebook Variable Name</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>What is the priority nursing intervention at this time?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a) Apply oxygen 4 L per rebreather mask</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) <strong>Administer 0.9% sodium chloride intravenously at 30 mL/kg</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Administer 10 units of regular insulin subcutaneously</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) Administer Vancomycin 1 gram intravenously</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rationale – Elevated serum lactate level (normal range for critical ill client less than 2 mmol/L). Client is in uncompensated respiratory acidosis with hypoxemia. Normal ABG ranges (Workman, 2018, p. 188):</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Partial pressure of oxygen (PaO₂) - 75 - 100 mmHg</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Partial pressure of carbon dioxide (PaCO₂) - 38 - 42 mmHg</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arterial blood pH of 7.38 - 7.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bicarbonate - (HCO₃⁻) - 22 - 28 mEq/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WBC is elevated and the presents of bands indicates a shift-to-the-left. H&amp;H and RBC levels are low normal values but would need to consider hemoconcentration effect which in essence would render these values low. Platelet level is low (normal range 150,000 – 450,000). H&amp;H, RBC, &amp; platelet compromise indicate disseminated intravascular coagulation. Serum glucose is elevated which is indicative of sepsis in the absence of diabetes (Heimgartner, 2018a, pp. 761-763).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Early intervention requires administration of fluids at 30 mL/kg within the first 3 hours of suspecting sepsis (Heimgartner, 2018a, p. 765). Oxygen would be administered but via a non-rebreather mask to decrease the amount of CO₂ the client rebreathes. Insulin may be given but fluid resuscitation is the priority. Antibiotics will be administered after blood cultures are obtained and fluid resuscitation has been provided (Heimgartner, 2018a, p. 765).</td>
</tr>
<tr>
<td>21</td>
<td>I_FW_E</td>
<td>Which of the following is an early indicator that the intervention the nurse performed is therapeutically effective?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a) Blood pressure 108/60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) WBC 8,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) <strong>Urine output 45 mL/hr</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) Arterial pH level 7.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rationale – return of adequate urinary output indicates improvement in fluid volume status (Heimgartner, 2018a, p. 765; Workman, 2018, p.</td>
</tr>
<tr>
<td>Test Item Number</td>
<td>Codebook Variable Name</td>
<td>Item</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------</td>
<td>------</td>
</tr>
<tr>
<td>165</td>
<td></td>
<td>BP reading is still hypotensive, WBC are low, and pH level continues to indicate acidosis.</td>
</tr>
</tbody>
</table>
| 25               | FE_FW_A                | Item #5 –  
Primary concept - fluid balance  
Secondary concept – perfusion  

The nurse assesses a client 2 days after diagnosis of ischemic cerebral vascular accident with left sided paralysis, aphasia and dysphagia. Client is receiving nothing by mouth (NPO) and has 0.9% sodium chloride (NaCl) at 125 mL/hr, and 2-in-1 parenteral nutrition (PN) at 95 mL/hr running through a peripherally inserted central catheter (PICC) line. Apical pulse 102, blood pressure 138/92, respirations 24 on 3 liters oxygen per nasal cannula. Fluid balance for the past 24 hours is positive 1200 mL. She has a history of chronic atrial fibrillation and is on telemetry.  

Which assessment should the nurse perform first?  

a) **Auscultate lungs**  
b) Neurological checks  
c) Calculate mean arterial blood pressure  
d) Analyze telemetry reading  

Rationale - Objective assessment cues indicating fluid overload include tachycardia, tachypnea, hypertension, and positive fluid balance for past 24 hours despite receiving IV fluids (Workman, 2018, p. 172). Subjective assessment findings include client has atrial fibrillation. The PN rate is greater than 60 mL/hr.  

Auscultation of lung sounds to determine if client is experiencing pulmonary edema from fluid overload is essential for early intervention. Assessment cues differentiating fluid overload problem from another ischemic cerebral vascular accident requires student to understanding increased osmotic effect of PN solution. |
| 26               | FE_FW_I                | Assessment findings:  
Bilateral auscultation of lungs reveals moist crackles in both lower lobes anteriorly and posteriorly. Pulse oximeter oxygen saturation reading is 88%. |
<table>
<thead>
<tr>
<th>Test Item Number</th>
<th>Codebook Variable Name</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Neurological checks are unchanged from the previous findings over the last 2 days. Capillary blood glucose level is 134 mg/dL. Client reports headache pain 3/10 on the nonverbal FACES pain rating scale. Telemetry reading indicates irregular rhythm, rate 98, atrial fibrillation. Auscultation of heart sounds reveals a grade 3/4 systolic murmur that was not present on previous assessments. Peripheral pulses are bounding 3+ and bilateral jugular vein distention is noted. What is the first thing the nurse should do prior to contacting the healthcare provider? a) Hold the parental nutrition infusion, and change the 0.9% infusion to 30 mL/hr b) Change the oxygen to 4 liters per nonrebreather face mask c) Administer furosemide 80 mg IV d) Insert a urinary catheter to facilitate accurate monitoring of output Rationale – The presence of moist crackles in both lower lobes indicates pulmonary edema. Combined with the low oxygenation saturation level, client is experiencing respiratory distress. Headache, heart murmur, bounding PP, and JVD are all signs of fluid overload (Willis &amp; Rebar, 2018, p. 1223; Workman, 2018, p. 172). All interventions presented in the answer options will need to be performed. The priority is to prevent the addition of high osmotic solutions from entering the vascular system and restore fluid balance. Healthcare provider should be notified for further orders. Close monitoring for rebound hypoglycemia should be conducted and if needed administration of D10W should be added to maintain normal serum glucose levels (Perry, Potter, Ostendorf, &amp; Laplante, 2018, p. 864; Willis &amp; Rebar, 2018, p. 1223-1224).</td>
</tr>
<tr>
<td>27</td>
<td>FE_FW_E</td>
<td>Which of the following diagnostic findings is an indicator that the intervention the nurse performed is achieving the desired therapeutic effect? a) Serum sodium level 128 mEq/L</td>
</tr>
<tr>
<td>Test Item Number</td>
<td>Codebook Variable Name</td>
<td>Item</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Serum potassium level 3.1 mEq/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Urine specific gravity 1.052</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>d) Serum osmolarity 290 mOsm/kg</strong></td>
</tr>
</tbody>
</table>

Rationale – Return of normal serum osmolarity (normal range 270-300 mOsm) indicates improvement in the fluid to solute balance (Workman, 2018, p. 173). All other values are abnormal – normal ranges (Workman, 2018, p. 164; Winkelman, 2018, p. 1333):
- Sodium 136 – 145 mEq/L
- Potassium 3.5 – 5.0 mEq/L
- Urine specific gravity 1.005-1.045
Appendix E

Clinical Judgment Ability Continuum

### Appendix F

**MCQ Test Items – Non-Framework Group**

<table>
<thead>
<tr>
<th>Test Item Number</th>
<th>Codebook Variable Name</th>
<th>Item</th>
</tr>
</thead>
</table>
| 4                | M_NFW_A                | An early symptom of hypoglycemia is?  
  a) **Irritability**  
  b) Warm moist skin  
  c) Ketones in urine  
  d) Extreme thirst |

Rationale - Decreased availability of glucose in cerebral circulation can cause change in level of consciousness. Early symptoms include anxiety, irritability and confusion. Warm moist skin, ketones in urine and extreme thirst are symptoms of hyperglycemia (Hendricks, 2018, p. 1309).

| 5                | M_NFW_I                | The client’s blood glucose level reading improves to 102 mg/dL after administration of 50 mL of D50W intravenous push for a blood glucose reading of 38 mg/dL. Which of the following interventions should the nurse perform next?  
  a) Notify the provider of the client’s hypoglycemic episode  
  b) Monitor the blood glucose readings every 15 minutes for the next hour  
  c) Teach the client early signs of hypoglycemia  
  d) **Provide the client with a high protein, complex carbohydrate snack** |

Rationale - After CBS returns to normal limits client should eat a small meal or snack containing complex carbohydrates and protein to provide slow steady resource of glucose absorption (Hendricks, 2018, p. 1310). 

The client should be taught early signs of hypoglycemia and how to prevent hypoglycemic episodes. Monitoring the CBS every 15 minutes is not necessary unless the client is in life threatening situation such as DKA. 

The provider may be informed of hypoglycemic episode but this is not the top priority at this time.

<p>| 6                | M_NFW_E                | While teaching a client who has a new diagnosis of diabetes mellitus, the nurse explains that to delay the onset of vascular complications it is important for the client to? |</p>
<table>
<thead>
<tr>
<th>Test Item Number</th>
<th>Codebook Variable Name</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>a) Exercise for at least 30 minutes every day</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) <strong>Maintain blood glucose levels within the normal range limits</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Avoid all stressful situations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) Eat a diet high in carbohydrates, low in protein and saturated fats</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Rationale</strong> – chronic hyperglycemia causes thickening to cell membranes, can be toxic to cell integrity and damages large and small blood vessels (Hendricks, 2018, p. 1283). Avoiding stressful situations and exercising 30 minutes every day are excellent health promotion activities but will not prevent cellular damage is hyperglycemic episodes occur. While no one dietary regimen is recommended for all client’s with diabetes, a diet of moderate complex carbohydrates, high in protein and low in saturated fats is recommended ((Hendricks, 2018, pp. 1299-1300).</td>
</tr>
<tr>
<td>10</td>
<td>P_NFW_A</td>
<td>Which of the following assessment findings is an early sign of a deep vein thrombosis?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a) 3+ edema to bilateral lower extremities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Large varicose veins to in both lower extremities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) <strong>Pain and swelling of left lower extremity</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) Shortness of breath on ambulation</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Rationale</strong> – classic signs of DVT are tenderness, pain, and sudden onset of unilateral swelling of the leg (Heimgartner, 2018b, pp. 742-743). Edema to bilateral lower extremities indicate decreased venous return. Large varicose vessels to bilateral lower extremities are not indicative of DVT. Short of breath with ambulation is also not an indicator of DVT.</td>
</tr>
<tr>
<td>11</td>
<td>P_NFW_I</td>
<td>The nurse notes bruising on a client’s upper extremities. Since the client is receiving warfarin the nurse will request an order for which laboratory test?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a) <strong>International normalized ratio (INR)</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Hemoglobin and hematocrit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Partial Thromboplastin time (PTT)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) Vitamin K level</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Rationale</strong> – INR is used to measure the therapeutic levels of warfarin (Heimgartner, 2018b, p. 745). Hemoglobin and hematocrit trending</td>
</tr>
<tr>
<td>Test Item Number</td>
<td>Codebook Variable Name</td>
<td>Item</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------</td>
<td>------</td>
</tr>
<tr>
<td>12</td>
<td>P_NFW_E</td>
<td>down may indicate a small bleed but is not used to determine therapeutic level of warfarin. PTT is used for heparin therapy. Vitamin K is the antidote for elevated warfarin levels (pp. 744-745). A client has been receiving heparin 5,000 units subcutaneously every 12 hours for 3 days. The RN determines the medication is the expected therapeutic effect by which of the following serum lab values?</td>
</tr>
</tbody>
</table>
|                  |                        | a) Hemoglobin 15 g/dL  
|                  |                        | b) RBC 4.8 m/μL  
|                  |                        | c) Platelet count 160,000/mm³  
|                  |                        | d) aPTT 1.5 times the normal control level |
| 16               | O_NFW_A                | Wheezing is noted during auscultation of the right upper lobe posteriorly. The nurse should? |
|                  |                        | a) Have the client cough and re-auscultate  
|                  |                        | b) Document the findings  
|                  |                        | c) Apply oxygen a 4 L per nonrebreather mask  
|                  |                        | d) Elevate the head of the bed to 45 degrees |
| 17               | O_NFW_I                | A client is receiving two medications by inhaler for the treatment of COPD. The nurse will administer which of the following medications first? |
|                  |                        | a) Beclomethasone  
|                  |                        | b) Albuterol  
<p>|                  |                        | c) Ipratropium |</p>
<table>
<thead>
<tr>
<th>Test Item Number</th>
<th>Codebook Variable Name</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>d) Montelukast</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rationale – Albuterol is a bronchodilator. When taking with other inhaled medications, the bronchodilator should be administered at least 5 minutes prior to other medications to promote bronchodilation effect to increase the penetration of the other inhaled medications (Workman, 2018, p. 569).</td>
</tr>
<tr>
<td>18</td>
<td>O_NFW_E</td>
<td>While teaching the client how to use an incentive spirometer the nurse determines the teaching has been effective if the nurse observes the client?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a) Place lips tightly around the mouthpiece and inhale slowly</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Place lips tightly around the mouthpiece and exhale slowly</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Place lips tightly around the mouthpiece and exhale as fast and as hard as possible</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) Place lips tightly around the mouthpiece and exhale until lungs are completely empty</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rationale – goal of incentive spirometer measurement is to determine the volume of air inhaled during deep breathing exercises to establish a baseline upon which to measure future assessments (Perry et al., 2018, pp. 639-642).</td>
</tr>
<tr>
<td>22</td>
<td>I_NFW_A</td>
<td>Which of the following may be a symptom of a urinary tract infection in an older adult?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a) Confusion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Slurred speech</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) 2+ edema to lower extremities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) Syncope</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rationale – confusion may be the only indicator of UTI in elderly clients as other symptoms are often missed (Camicia &amp; Ignatavicius, 2018, p. 91). All other answer choices are not indicative of UTI.</td>
</tr>
<tr>
<td>23</td>
<td>I_NFW_I</td>
<td>Which of the following techniques is required to prevent central venous access device associated infections?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a) Clean the injection hubs with antiseptic solution for 15 seconds and allow to dry prior to using</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Change the central venous access device dressing every day</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Change the intravenous tubing every shift</td>
</tr>
<tr>
<td>Test Item Number</td>
<td>Codebook Variable Name</td>
<td>Item</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) <strong>Always apply sterile gloves prior to administering medications through a central venous access device</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rationale – (Adams, 2016, p. 216; Perry et al., 2018, p. 792). Dressing changes are not required daily. Tape and gauze dressings should be changed every 48 hours, transparent membrane dressings every 5 to 7 days (Adams, 2018, p. 213). Once tubing is in place, if the closed system is maintained tubing should be change every 4 to 7 days to limit the number of times the system is opened (p. 214). Hand hygiene with clean gloves should be used for medication administration. Sterile gloves are not required (Perry et al., 2018, p. 610).</td>
</tr>
<tr>
<td>24</td>
<td>I_NFW_E</td>
<td>The nurse will contact the health care provider <em>prior</em> to administering vancomycin 1 gram IVPB to determine if the medication should be given for which of the following?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a) The serum trough level is low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) <strong>The serum trough level is high</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) The serum peak level is low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) The serum peak level is high</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rationale – Maintaining therapeutic concentration range of vancomycin in the plasma is essential to achieve desired effect as well as prevent drug toxicity (Kee, Hayes, &amp; McCuistion, 2015, p. 10 &amp; 401). The trough level is drawn within 30 minutes prior to administration of the medication. The peak level is drawn approximately 1 hour after administration.</td>
</tr>
<tr>
<td>28</td>
<td>FE_NFW_A</td>
<td>Which of the following findings would be of most concern for a client receiving furosemide?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a) Increased urinary output</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Dizziness upon standing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) <strong>Muscle weakness and cramps in lower extremities and hands</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) Decrease in blood pressure from 168/92 to 128/78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rationale – furosemide is a loop-diuretic that inhibit sodium reabsorption in the kidney tubules. A frequent complication is hypokalemia. Muscle weakness and cramps are common symptoms of low serum potassium levels (Kee et al., 2015, pp. 624-625; Workman, 2018, pp. 175-176). Increased urinary output and a decrease in blood pressure are expected therapeutic outcomes of</td>
</tr>
<tr>
<td>Test Item Number</td>
<td>Codebook Variable Name</td>
<td>Item</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>diuretic therapy. Dizziness upon standing indicates orthostatic hypotension that can occur until the body adjust to lower BP state. Fall precautions and patient education should be provided.</td>
</tr>
<tr>
<td>29</td>
<td>FE_NFW_I</td>
<td>Which of the following intravenous fluid orders would be most appropriate for a client with heart failure?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a) D5 and 0.9% Sodium Chloride at 185 mL/hr</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) <strong>D5W at 30 mL/hr</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) LR and 0.45% Sodium Chloride at 75 mL/hr</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) 0.9% Sodium Chloride at 50 mL/hr</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rationale – D5W is an isotonic solution that does not contain sodium. Items a &amp; b are hypertonic solutions and contain sodium. Item d is isotonic but also contains sodium (Workman, 2018, p. 170). Client’s with heart failure are restricted to low sodium diets (usually 2-3 grams per day) to prevent water retention and fluid restriction of 2 L per day (Dechant, 2018, p. 704). D5W at 30 mL/hr is the only answer selection that meets these criteria.</td>
</tr>
<tr>
<td>30</td>
<td>FE_NFW_E</td>
<td>Which of the following assessment findings indicates the diuretic medication is achieving the therapeutic effect?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a) Heart rate increases from 58 to 72 beats per minute</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Blood pressure increases from 104/62 to 128/72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c) Weight decreases from 87 kg to 77 kg over one month</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d) <strong>Urinary output increases from 30 mL/hr to 60 mL/hr</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rationale – Diuretic medications promote diuresis by inhibiting sodium and water reabsorption form the kidney tubules, increasing urinary output (Kee et al., 2015, pp. 620-622). With decreased vascular volume blood pressure decreases. Increase in pulse rate would indicate compensatory mechanism for fluid volume deficit. Weight may decrease with fluid loss but is not the primary indicator of therapeutic effectiveness.</td>
</tr>
</tbody>
</table>
# Appendix G

## Test Blueprint

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>M_FW_A</td>
<td>Metabolism</td>
<td>Reduction in Risk Potential</td>
<td>Assess</td>
<td>Perfusion</td>
<td>Cure Recognition</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>M_FW_I</td>
<td>Metabolism</td>
<td>Reduction in Risk Potential</td>
<td>Implement</td>
<td></td>
<td>Take Action</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>M_FW_E</td>
<td>Metabolism</td>
<td>Physiological Adaptation</td>
<td>Evaluate</td>
<td></td>
<td>Evaluate Outcomes</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>M_NF_W_A</td>
<td>Metabolism</td>
<td>Reduction in Risk Potential</td>
<td>Assess</td>
<td></td>
<td>Remember/Understand</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>M_NF_W_I</td>
<td>Metabolism</td>
<td>Health Promotion and Maintenance</td>
<td>Implement</td>
<td></td>
<td>Analyze/Apply</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>M_NF_W_E</td>
<td>Metabolism</td>
<td>Reduction in Risk Potential</td>
<td>Evaluate</td>
<td></td>
<td>Evaluate</td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>P_FW_A</td>
<td>Perfusion</td>
<td>Reduction in Risk Potential</td>
<td>Assess</td>
<td>Metabolism</td>
<td>Cue Recognition</td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>P_FW_I</td>
<td>Perfusion</td>
<td>Reduction in Risk Potential</td>
<td>Implement</td>
<td></td>
<td>Take Action</td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>P_FW_E</td>
<td>Perfusion</td>
<td>Physiological Adaptation</td>
<td>Evaluate</td>
<td></td>
<td>Evaluate Outcomes</td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>P_NF_W_A</td>
<td>Perfusion</td>
<td>Pharmacological and Parenteral Therapies</td>
<td>Assess</td>
<td></td>
<td>Remember/Understand</td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>P_NF_W_I</td>
<td>Perfusion</td>
<td>Health Promotion and Maintenance</td>
<td>Implement</td>
<td></td>
<td>Analyze/Apply</td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>P_NF_W_E</td>
<td>Perfusion</td>
<td>Reduction in Risk Potential</td>
<td>Evaluate</td>
<td></td>
<td>Evaluate</td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td>O_FW_A</td>
<td>Oxygenation</td>
<td>Reduction in Risk Potential</td>
<td>Assess</td>
<td>Fluid &amp; Electrolytes</td>
<td>Cue Recognition</td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td>O_FW_I</td>
<td>Oxygenation</td>
<td>Reduction in Risk Potential</td>
<td>Implement</td>
<td></td>
<td>Take Action</td>
<td></td>
</tr>
<tr>
<td>15.</td>
<td>O_FW_E</td>
<td>Oxygenation</td>
<td>Physiological Adaptation</td>
<td>Evaluate</td>
<td></td>
<td>Evaluate Outcomes</td>
<td></td>
</tr>
<tr>
<td>16.</td>
<td>O_NF_W_A</td>
<td>Oxygenation</td>
<td>Physiological Adaptation</td>
<td>Assess</td>
<td></td>
<td>Remember/Understand</td>
<td></td>
</tr>
<tr>
<td>17.</td>
<td>O_NF_W_I</td>
<td>Oxygenation</td>
<td>Health Promotion</td>
<td>Implement</td>
<td></td>
<td>Analyze/Apply</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------</td>
<td>-----------------------</td>
<td>-------------------------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>-------------------------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>18</td>
<td>O_NF W_E</td>
<td>Oxygenation</td>
<td>Reduction in Risk Potential</td>
<td>Evaluate</td>
<td></td>
<td></td>
<td>Evaluate</td>
</tr>
<tr>
<td>19</td>
<td>I_FW_A</td>
<td>Infection</td>
<td>Reduction in Risk Potential</td>
<td>Assess</td>
<td>Oxgenation</td>
<td></td>
<td>Cue Recognition</td>
</tr>
<tr>
<td>20</td>
<td>I_FW_I</td>
<td>Infection</td>
<td>Safety and Infection Control</td>
<td>Implement</td>
<td></td>
<td></td>
<td>Take Action</td>
</tr>
<tr>
<td>21</td>
<td>I_FW_E</td>
<td>Infection</td>
<td>Physiological Adaptation</td>
<td>Evaluate</td>
<td></td>
<td></td>
<td>Evaluate Outcomes</td>
</tr>
<tr>
<td>22</td>
<td>I_NF W_A</td>
<td>Infection</td>
<td>Pharmacological and Parenteral Therapies</td>
<td>Assess</td>
<td></td>
<td></td>
<td>Remember/Understand</td>
</tr>
<tr>
<td>23</td>
<td>I_NF W_I</td>
<td>Infection</td>
<td>Health Promotion and Maintenance</td>
<td>Implement</td>
<td></td>
<td></td>
<td>Analyze/Apply</td>
</tr>
<tr>
<td>24</td>
<td>I_NF W_E</td>
<td>Infection</td>
<td>Reduction in Risk Potential</td>
<td>Evaluate</td>
<td></td>
<td></td>
<td>Evaluate</td>
</tr>
<tr>
<td>25</td>
<td>F_E_F W_A</td>
<td>Fluid and Electrolytes</td>
<td>Reduction in Risk Potential</td>
<td>Assess</td>
<td>Perfusion</td>
<td></td>
<td>Cue Recognition</td>
</tr>
<tr>
<td>26</td>
<td>F_E_F W_I</td>
<td>Fluid and Electrolytes</td>
<td>Pharmacological and Parenteral Therapies</td>
<td>Implement</td>
<td></td>
<td></td>
<td>Take Action</td>
</tr>
<tr>
<td>27</td>
<td>F_E_F W_E</td>
<td>Fluid and Electrolytes</td>
<td>Physiological Adaptation</td>
<td>Evaluate</td>
<td></td>
<td></td>
<td>Evaluate Outcomes</td>
</tr>
<tr>
<td>28</td>
<td>F_E_NF W_A</td>
<td>Fluid and Electrolytes</td>
<td>Pharmacological and Parenteral Therapies</td>
<td>Assess</td>
<td></td>
<td></td>
<td>Remember/Understand</td>
</tr>
<tr>
<td>29</td>
<td>F_E_NF W_I</td>
<td>Fluid and Electrolytes</td>
<td>Pharmacological and Parenteral Therapies</td>
<td>Implement</td>
<td></td>
<td></td>
<td>Analyze/Apply</td>
</tr>
<tr>
<td>30</td>
<td>F_E_NF W_E</td>
<td>Fluid and Electrolytes</td>
<td>Reduction in Risk Potential</td>
<td>Evaluate</td>
<td></td>
<td></td>
<td>Evaluate</td>
</tr>
</tbody>
</table>
Appendix H

Informed Consent

Date: May 2019
Project Title: The Effectiveness of Multiple-Choice Test Items in Assessing the Clinical Judgment Abilities of Prelicensure Registered Nursing Students
Principal Investigator: Tracy Holt
Capstone Advisor: Dr. Sandra Cleveland
Version #: 1
Approval Date: April 12, 2019
Approved Consent is valid for one year from the date of IRB approval.

You are being asked to participate in a DNP capstone project. This form provides you with information about the project. The project will be described all of your questions will be answered before you sign this consent. Please read the information below and ask questions about anything you do not understand before deciding whether or not to take part in this project.

Why is this project being done?
The purpose of this project is to explore the effectiveness of multiple-choice test items in assessing the clinical judgment abilities of registered nursing students in prelicensure programs.

You are being asked to take part in this project because you are a registered nursing student in an associate degree program. Approximately 70 students will take part in this project.

What happens if I participate in this project?
If you agree to take part in this project, you will be asked to complete an online examination consisting of 30 multiple-choice test items. You will be provided test student user login and passwords to keep your participation strictly anonymous. In the event of technical difficulties, a written version of the examination will be provided.

Your participation will last not more one hour.

What are the possible discomforts or risks?
Discomforts you may experience while taking part in this project include the time it will take you to answer all the questions on the exam. Other possible discomforts or risks may include stress while completing the exam. If you experience any emotional distress as a result of taking part in this project and would like to talk with someone at your own expense, you can access student services located in building 7 (Phone: [blank space for phone number]).

Subject Initials
What are the possible benefits of the project?

This project/study is designed to learn more about the new type of test items on the Next Generation NCLEX-RN® examination. There are no direct benefits to you for participating in this project/study. However, learning more about the new type of test items on the NCLEX-RN® examination may help you prepare to study for your upcoming licensing examination.

Who is paying for this project? There is no funding for this project.

Will I be paid for being in the study? Will I have to pay for anything?

You will not be paid to participate in the project and it will not cost you anything to participate in this project. A certificate of participation for your professional portfolio will be provided to all participants.

Is my participation voluntary?

Taking part in this project is voluntary. You have the right to choose not to take part in this project. If you choose to take part, you have the right to stop at any time. If you refuse or decide to withdraw later, you will not lose any benefits or rights to which you are entitled.

Who do I call if I have questions?

The principal investigator (student) carrying out this project is Tracy Holt. If you have questions, you may contact the principal investigator by emailing [email protected]. You may have questions about your rights as someone in this study. You can contact the students’ Capstone advisor, Dr. Sandra Cleveland at [email protected] or by sending a message to [email protected]. You may also contact the American Sentinel University IRB Director or the Associate Dean of Graduate programs with questions about your rights as a research subject at [email protected]

Who will see my information?

A test student login user ID and password will be provided to you to participate in the online examination. Your name will not be required.

If technical difficulties prevent the use of the computers for the exam, the written copies of the exam as well as the answer sheets will be numbered by the researcher. Your name will not be required. The only document that will require personal identification data is this consent form. An envelope will be provided for you to seal your signed consent form immediately upon completion. The envelopes will be placed in a locked drawer in the researcher’s home filing cabinet. The forms will be destroyed after five years.

Any information you provide during this project will be kept strictly private (confidential). Any information you provide may be looked at by the following:

- The DNP students’ Capstone chair and committee members
- American Sentinel University Institutional Review Board (IRB)
- Regulatory officials from the institution where the project is being conducted who want to make sure the research is safe

Subject Initials
The results from this project may be shared at a meeting with the DNP students’ capstone committee, at a professional conference, and may also be in published articles. Your name will be kept private when information about this project is presented in any form.

**Agreement to be in this study/project**

I have read this paper about the project or it was read to me. I understand the possible risks and benefits of this study. I know that taking part in this project is voluntary. I choose to take part in this study/ and I will get a copy of this consent form.

Subject - Signature: ____________________________________ Date: _________________

Subject - Print Name: __________________________________________________________

Consent form explained by: ______________________________ Date: _________________

Print Name: ______________________________
Appendix I

Proctor Script and Administration Instructions

Thank you for proctoring this examination. The students were previously provided with a copy of the consent form explaining the research process. Please re-emphasize that all participation is voluntary. Have students remove all items from their desk except personal items such as water or tissues. Ask if there are any questions prior to providing the access code. Once the examination has begun, do not answer student’s questions or provide any further information regarding the test questions or answer choices. You may provide assistance as needed with computer technical difficulties.

The first screen contains the consent form and must be acknowledged indicating informed consent is granted to access the examination. Each exam item will be presented one at a time. Each question must be answered and submitted prior to moving to the next item. Back tracking will not be allowed. The final screen contains items for collection of demographical data and should be completed prior to submitting the exam for grading. Access per participant is set at one-time only to prevent resubmission. Once the exam is submitted for grading, the correct answers and rationales will appear for independent review. Please ask the students to remain quietly seated to prevent distracting other students who are still taking the exam. After all participants have completed the examination please conduct a review of the examination with all participants collectively to clarify any items to prevent confusion or misunderstandings.

Thank you,

Tracy D. Holt - DNP Nursing Educational Leadership Student
American Sentinel University

Electronic copy available at: https://ssrn.com/abstract=3634804
Appendix J

Placeholder for letter of approval.
Appendix K

Theoretical Framework: The Nursing Process with Corresponding Processes of Clinical Judgment

(Doenges & Moorhouse, 2013; Hooper, 2017, November; Muntean et al., 2015)
Appendix M

IRB Approval from American Sentinel University

April 12, 2019

Tracy Holt
DNP Student
American Sentinel University

Re: The Effectiveness of Multiple-Choice Test Items in Assessing the Clinical Judgment Abilities of Premedication Registered Nursing Students

Dear Ms Holt,

On April 12, the Institutional Review Board (IRB) of American Sentinel University has approved the above-referenced submission and has deemed it as an expedited study. The contingencies have been addressed and the IRB approves the protocol. Work on this project may begin. This approval is for a period of one year from the dates of this letter and will require continuation approval if the research extends beyond one year. If you make changes to the protocol during the period of this approval, you must submit a revised protocol to the American Sentinel University IRB for approval before implementing the changes.

If you have any questions regarding the IRB’s decision, please contact me through

Sincerely,

[Name redacted]
Elaine Foster PhD, MSN, RN
Chair
American Sentinel University IRB

C. Dr. Cleveland
Appendix N

IRB Approval from College
NCLEX-RN EXAM PRACTICE

WHEN
Thursday May 16th
3 pm to 4 pm

WHERE
Room 13-342

FEATURING - New NCLEX-RN Style Questions

CALLING ALL RN STUDENTS

VOLUNTEERS NEEDED
- This event is not associated with the Nursing Program
- Practice exam scores will NOT be included in your course grade

BENEFITS
- NCLEX-RN practice exam
- Discussion of rationales
- Explanation of new test item types
- Certificate of participation for portfolio

For more information email Tracy Hilt DNP student at
Appendix P

Permission to Use NCSBN ® Nursing Clinical Judgement Model

3/5/2019

Gmail - Approval for use of NCSBN Nursing Clinical Judgement Model For DNP Research Proposal

TH [Redacted]

Approval for use of NCSBN Nursing Clinical Judgement Model For DNP Research Proposal

1 message

Philip Dickison [Redacted]

To: [Redacted]

Tue, Mar 5, 2019 at 10:03 AM

Dear Ms. Holt:

NCSBN is pleased to provide you with approval to use the NCSBN Nursing Clinical Judgment Model with appropriate reference for your research project Effectiveness of Multiple-Choice Test Items in Assessing the Clinical Judgment Abilities of Prelicensure Registered Nursing Students. We would be interested in the findings of your research once it is completed.

Regards

Philip D. Dickison, PhD, RN
Chief Officer, Operations and Examinations
National Council State Boards of Nursing

The information contained in this e-mail and any accompanying documents is intended for the sole use of the recipient to whom it is addressed, and may contain information that is privileged, confidential, and prohibited from disclosure under applicable law. If you are not the intended recipient, or authorized to receive this on behalf of the recipient, you are hereby notified that any review, use, disclosure, copying, or distribution is prohibited. If you are not the intended recipient(s), please contact the sender by e-mail and destroy all copies of the original message. The integrity and security of this message cannot be guaranteed on the Internet. Thank you.
Appendix Q

Email Invitation to Students

Dear RN Students,

I am looking for volunteers to help me conduct a study for my DNP Capstone Project. As you most probably are aware, the NCLEX-RN® is changing. My project consists of test items written using the new NCLEX-RN® format.

Your decision to participate or not participate will not have any impact on your current nursing courses or your nursing program. Participation is strictly voluntary. No personal information will be collected. Your participation will be anonymous. Those who do participate will receive a certificate of appreciation for your participation that you can include in your professional portfolio.

If you are interested, here are the details:

- The study consists of the completion of a 30-item examination using Canvas® administered on May 16th at 3 pm in the nursing computer lab
- The exam will be open for 60 minutes and you may leave at any time
- The correct answers will be provided, and the rationales explained after everyone has completed the exam

No need to email or sign-up, just show up in the computer room at the scheduled date and time.

Sincerely,
Tracy Holt
American Sentinel University
DNP student
Appendix R

Certificate of Participation

Certificate of Appreciation

Thank you for participating in the Capstone Research Project
by
Tracy Holt - DNP Student
Spring, 2019
Appendix S

Project Codebook

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Variable Type</th>
<th>Label</th>
<th>Values</th>
<th>Level of Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>StudentID</td>
<td>String</td>
<td>Student ID</td>
<td>Alphanumeric</td>
<td>N/A</td>
</tr>
<tr>
<td>Age_Yrs</td>
<td>Numeric-Continuous</td>
<td>Age in Years</td>
<td>18-25 = 1; 26-30 = 2; 31-35 = 3; 36-40 = 4; &gt;40 = 5</td>
<td>Ordinal</td>
</tr>
<tr>
<td>Gender</td>
<td>Numeric-Categorical</td>
<td>Gender</td>
<td>Male=1; Female=2; Other=3</td>
<td>Nominal</td>
</tr>
<tr>
<td>Race</td>
<td>Numeric-Categorical</td>
<td>Race</td>
<td>American Indian = 1; Asian = 2; African American = 3; Caucasian = 4; Other = 5</td>
<td>Nominal</td>
</tr>
<tr>
<td>GPA</td>
<td>Numeric-Continuous</td>
<td>Overall GPA</td>
<td>2.2-2.4 = 1; 2.5-3.0 = 2; 3.1-3.4 = 3; 3.5-4.0 = 4</td>
<td>Ordinal</td>
</tr>
<tr>
<td>Degree</td>
<td>Numeric-Categorical</td>
<td>College degree or certificate – No, Yes, non-healthcare, Yes, healthcare</td>
<td>No=0; Yes=1; Yes=2</td>
<td>Nominal</td>
</tr>
<tr>
<td>Emp</td>
<td>Numeric-Categorical</td>
<td>Work experience - No, Yes, non-healthcare, Yes, healthcare</td>
<td>No=0; Yes=1; Yes=2</td>
<td>Nominal</td>
</tr>
<tr>
<td>M_FW_A</td>
<td>Numeric-Categorical</td>
<td>Metabolism – cue recognition</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td>M_FW_I</td>
<td>Numeric-Categorical</td>
<td>Metabolism – take action</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td>M_FW_E</td>
<td>Numeric-Categorical</td>
<td>Metabolism – evaluate outcomes</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td>M_NFW_A</td>
<td>Numeric-Categorical</td>
<td>Metabolism – remember/understand</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td></td>
<td>Numeric-Categorical</td>
<td>Metabolism – analyze/apply</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td>---</td>
<td>-------------------</td>
<td>--------------------------</td>
<td>------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>M_NFW_I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M_NFW_E</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P_FW_A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P_FW_I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P_FW_E</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P_NFW_A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P_NFW_I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P_NFW_E</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O_FW_A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O_FW_I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O_FW_E</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O_NFW_A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O_NFW_I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O_NFW_E</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I_FW_A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I_FW_I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I_FW_E</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>Data Type</td>
<td>Description</td>
<td>Score</td>
<td>Scale</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>---------------------------------------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>I_NFW_A</td>
<td>Numeric-Categorical</td>
<td>Infection – remember/understand</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td>I_NFW_I</td>
<td>Numeric-Categorical</td>
<td>Infection – analyze/apply</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td>I_NFW_E</td>
<td>Numeric-Categorical</td>
<td>Infection – evaluate</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td>FE_FWA</td>
<td>Numeric-Categorical</td>
<td>Fluid/electrolytes – cue recognition</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td>FE_FWI</td>
<td>Numeric-Categorical</td>
<td>Fluid/electrolytes – take action</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td>FE_FWE</td>
<td>Numeric-Categorical</td>
<td>Fluid/electrolytes – evaluate outcomes</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td>FE_NFW_A</td>
<td>Numeric-Categorical</td>
<td>Fluid/electrolytes – remember/understand</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td>FE_NFW_I</td>
<td>Numeric-Categorical</td>
<td>Fluid/electrolytes – analyze/apply</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td>FE_NFW_E</td>
<td>Numeric-Categorical</td>
<td>Fluid/electrolytes – evaluate</td>
<td>Incorrect=0; Correct=1</td>
<td>Nominal</td>
</tr>
<tr>
<td>Tot_FW</td>
<td>Numeric-Continuous</td>
<td>Total framework questions correct</td>
<td>0 to 1050</td>
<td>Scale</td>
</tr>
<tr>
<td>Tot_NFW</td>
<td>Numeric-Continuous</td>
<td>Total non-framework questions correct</td>
<td>0 to 1050</td>
<td>Scale</td>
</tr>
<tr>
<td>Tot_FW_A</td>
<td>Numeric-Continuous</td>
<td>Total framework assessment questions correct</td>
<td>0 to 350</td>
<td>Scale</td>
</tr>
<tr>
<td>Tot_FW_I</td>
<td>Numeric-Continuous</td>
<td>Total framework implementation questions correct</td>
<td>0 to 350</td>
<td>Scale</td>
</tr>
<tr>
<td>Tot_FW_E</td>
<td>Numeric-Continuous</td>
<td>Total framework evaluation questions correct</td>
<td>0 to 350</td>
<td>Scale</td>
</tr>
<tr>
<td>Tot_NFW_A</td>
<td>Numeric-Continuous</td>
<td>Total non-framework assessment questions correct</td>
<td>0 to 350</td>
<td>Scale</td>
</tr>
<tr>
<td>Tot_NFW_I</td>
<td>Numeric-Continuous</td>
<td>Total non-framework implementation questions correct</td>
<td>0 to 350</td>
<td>Scale</td>
</tr>
<tr>
<td>Variable</td>
<td>Type</td>
<td>Description</td>
<td>Range</td>
<td>Scale</td>
</tr>
<tr>
<td>------------</td>
<td>------------</td>
<td>-------------------------------------------------------</td>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>Tot_NFW_E</td>
<td>Numeric-continuous</td>
<td>Total non-framework evaluation questions correct</td>
<td>0 to 350</td>
<td>Scale</td>
</tr>
<tr>
<td>FW_CJ</td>
<td>Categorical</td>
<td>Individual framework item answered correctly or not</td>
<td>0=Incorrect; 1=Correct</td>
<td>Nominal</td>
</tr>
<tr>
<td>NFW_CJ</td>
<td>Categorical</td>
<td>Individual non-framework item answered correctly or not</td>
<td>0=Incorrect; 1=Correct</td>
<td>Nominal</td>
</tr>
</tbody>
</table>