Assessing Clinical Reasoning Competency in Enrolled Nursing Students: Development and Validation of a Novel Clinical Competency Assessment Tool

by

LEUNG, Lai Ching Christee

A Thesis Submitted to

The Education University of Hong Kong

in Partial Fulfillment of the Requirement for

the Degree of Doctor of Education

April 2020

ProQuest Number: 28150025

ProQuest 28150025

Published by ProQuest LLC (2020). Copyright of the Dissertation is held by the Author.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Statement of Originality

I, LEUNG, Lai Ching Christee, hereby declare that I am the sole author of the thesis and the material presented in this thesis is my original work except those indicated in the acknowledgement. I further declare that I have followed the University's policies and regulations on Academic Honesty, Copyright and Plagiarism in writing the thesis and no material in this thesis has been submitted for a degree in this or other universities.

Abstract

Background: As the rapid growth of the diversity and complexity of clinical situations in Hong Kong, nurses always encounter different clinical situations of uncertainty in daily practice that need to make timely decisions and implement immediate nursing interventions, therefore, clinical reasoning competency is a crucial capability that the nurses are required to have; however, there are lack of assessment tools designed for measuring the clinical reasoning competency of nursing students especially those who are in transition to the qualified nurses; thus, construction of a clinical reasoning competency assessment tool (CRCAT) is necessary.

Aim: To establish a reliable and valid clinical assessment tool, named "Clinical Reasoning Competency Assessment Tool (CRCAT)", to examine the clinical reasoning competency of the Enrolled Nursing students.

Design: A cross-sectional design

Participants: Forty-one Enrolled Nursing students from first year class and Forty-one from final year class at the school of general nursing training in Hong Kong and fifteen clinical nurses who are the clinical nursing instructors of the Enrolled Nursing students. Ninety-seven participants in total.

Methods: This study has two phases. Phase one is the development of clinical reasoning competency assessment tool (CRCAT). Phase two is the establishment of validity and reliability of CRCAT. All the participants including the Enrolled Nursing students and the clinical nursing instructors completed the CRCAT and the problem solving inventory (PSI) within the allowed time. The Enrolled Nursing students in first year class repeated the CRCAT two weeks after the first attempt.

Results: The results show that the clinical reasoning competency assessment tool (CRCAT) has good model-data fit with high intra-class correlation coefficient at 0.829, the person reliability with Cronbach alpha coefficient is acceptable at 0.64. ANOVA shows a significant difference in clinical reasoning competency among the three groups ($p < 0.05$). A significant difference ($p = 0.01$) in clinical reasoning competency found between the two groups of Enrolled Nursing students.

Conclusion: The clinical reasoning competency assessment tool (CRCAT) can be used as a clinical assessment tool to assess the clinical reasoning competency for the Enrolled Nursing students. Further study using a larger sample size with extended scope of sample subjects to improve the representative and generalizability of the CRCAT is necessary.

Acknowledgments

There are many individuals to whom I am indebted for their contribution to the completion of this research project. They are my constant source of inspiration and energy.

Firstly, I would like to give my wholehearted thanks to my supervisors, Prof. W. C. Wang, Prof. Joanne W. Y. Chung, Dr. Peggy P. L. Or and Dr. X. L. Qiu for their patience, motivation, enthusiasm and immerse knowledge. Their genuine support, guidance, advice and encouragement throughout the proposal stage to the thesis writing stage are indeed the determinants of the success of this study.

Secondly, I would like to express my appreciation to Ms. S. M. Lee and Ms. S. F. Yeung for their valuable time, opinions and suggestions on the content validity of the clinical reasoning competency assessment tool (CRCAT) developed in this study.

My sincere thanks also go to Prof. Puncky Heppner for his generous support, response to my requests and provision of a copy of Problem Solving Inventory (PSI) used in his previous study.

Last but not the least, I would like to thank my mother, brother and sisters for their continuing support and encouragement throughout my study.

Table of Contents

Statement of Originality	ii
Abstract	iii
Acknowledgements	v
Table of Contents	iv
List of Abbreviations	vii
List of Figures	viii
List of Tables	ix
Chapter 1: Introduction	1
Chapter 2: Literature Review	9
Chapter 3: Study Methodology	49
Chapter 4: Study Results of Phase One	57
Chapter 5: Study Results of Phase Two	74
Chapter 6: Discussion and Implication	94
Chapter 7: Conclusion	120
References	124
Appendix A: Clinical Reasoning Competency Assessment Tool (CRCAT)	135
Appendix B: Problem Solving Inventory (PSI)	146

List of Abbreviations

CRCAT	Clinical Reasoning Competency Assessment Tool
PSI	Problem Solving Inventory

List of Figures

Figure 2-1	Outcome Present State Test (OPT) Model	14
Figure 2-2	Reflective Self-regulated Learning Model	15
Figure 2-3	The clinical reasoning cycle	18
Figure 2-4	The six levels of the cognitive domain of Bloom's original taxonomy	31
Figure 2-5	Bloom's original and revised taxonomies	34
Figure 4-1	Conceptual framework of four building blocks	58
Figure 4-2	A Construct Map in Construct "Clinical Reasoning Competency"	61
Figure 5-1	Item Pathway Map	86
Figure 5-2	Item-person map	88

List of Tables

Table 2-1	Revised Bloom's Taxonomy Action Verbs	36
Table 2-2	Revised Bloom's Knowledge Dimension	37
Table 4-1	Scale descriptions and competency levels of the 5-pts Likert scale with scenario one	66
Table 4-2	Illustration of five competency levels of clinical reasoning with scenario one	68
Table 4-3	Understandability of the CRCAT questionnaire	72
Table 5-1	Demographics of nursing students	75
Table 5-2	Demographics of clinical nursing tutors	77
Table 5-3	Correlations between demographics and clinical reasoning competency	78
Table 5-4	Test Re-test Reliability of CRCAT	79
Table 5-5	Summary of 97 measured persons	81
Table 5-6	Summary of 60 measured items	81
Table 5-7	Item Fit Order	83
Table 5-8	Items Correlation Order	85
Table 5-9	Comparison among three groups using ANOVA	90
Table 5-10	Descriptive of clinical reasoning competency of three groups	91
Table 5-11	Multiple Comparisons of clinical reasoning competency among three groups	91

CHAPTER ONE

INTRODUCTION

1.1 Background of the study

The medical incident rate at the Hong Kong hospitals has been increasing in recent years. Some of the medical incidents reported in daily newspapers had revealed the nursing students' clinical performance. In one medical incident, an elderly patient died after a nursing procedure of tracheostomy suction performed by a nursing student, the investigation team evaluated that the improper position and insecure of the tracheostomy tube with tapes was one of the fatal causes. In another medical incident, a baby was severely scalded during the baby bathing procedure, the investigation team evaluated that the nursing student used elbow skin instead of water thermometer to test the temperature of bathing water was the main cause of this incident. In fact, almost all the reported medical incidents involving nursing are the common nursing procedures, which are to be assessed in the final clinical examination before graduation. Therefore, the medical incidents relating to nursing certainly quicken the nursing professionals consciously to review, analyze and evaluate the clinical practice in nursing with a view to gaining insights to improve the capability of clinical nursing performance.

1.2 Justification for the study

1.2.1 Nursing profession and competence

Professional competence is the required professional standard that is often specified by law that raises a professional in fulfilling the role particularly in his or her profession. The professional competence has several dimensions, such as pedagogical competence includes teacher's competence regarding collaboration, comprehensive view and contribution to the development of pedagogy for higher education (Suciu & Mata, 2011). In fact, according to the Nursing Council of Hong Kong, there are four areas of competence that the final year Enrolled Nursing students especially those who are in transition to enrolled nurses are required to reach for the licensed recognition including competence area 1: Professional, Legal and Ethical Nursing Practice; competence area 2: Provision of Care; competence area 3: Personal and Professional Attribute; and competence area 4: Teamwork. According to the professional standard as required by the Nursing Council of Hong Kong, the Enrolled Nurses are expected to have the ability to provide quality client-oriented care as well as to contribute to quality and risk management. In other words, the nursing students who are going into the transitional period to the qualified nurses are needed to be well equipped with these competence areas in order to be eligible for the licensed practicing certificate.

1.2.2 Clinical reasoning competency in nursing

The Competence Outcome Performance Assessment (COPA) Model was a comprehensive core competencies method designed model used to promote competence-focused and practice-focused in nursing education. In the model, all nursing skills and abilities were clustered under eight core practice competence categories including assessment and intervention, communication, critical thinking, human caring relationships, teaching, management, leadership, and knowledge integration skills (Armstrong, Spencer and Lenburg, 2009). Hence, competence in nursing is not focusing on one area but consisting of several core competencies. Critical thinking was one of the core practice competencies. In fact, there are multiple concepts have been used synonymously in the literature including decision-making, problem-solving, critical thinking, clinical judgment, diagnostic reasoning and clinical reasoning, while these concepts all include elements of both process and outcome focusing on the thinking strategies that a nurse uses to make a judgment or decision, and solve problem (Murphy 2004, Kautz et al. 2005, Su et al. 2005, Simmons, B. 2010). Clinical reasoning competency is one of the core clinical competencies that are required in nursing practice.

As the rapid growth of the diversity and complexity of clinical situations in Hong Kong, and the frequent change of new and advanced medical clinical technology, the needs for the health care team with good professional standards as well as the public expectations become increasingly high and demanding. Nurses, especially those who are working in the frontline, may encounter different clinical situations of uncertainty in daily practice that need to make timely decisions and implement immediate nursing interventions. Therefore, clinical reasoning is a crucial competency that the nurses are required to have. Certainly, the nursing students are expected to be well equipped. In fact, various pedagogical methods have already been integrated into the nursing curriculum in order to develop and enhance nursing students' learning in the area of clinical reasoning during their training years; however, according to the research finding of Hunter and Arthur (2016), the clinical educators comments on the current clinical performance assessment tool were consistent that the tool was unable to adequately appraise students' clinical reasoning ability during clinical placement. Moreover, Wang (2010) pointed out that the judgment of person ability level with raw score was test dependent; in other words, whether or not the results of the current assessments, which were raw score calculation, used in the hospital-based nursing schools under the Hospital Authority in Hong Kong reflected the nursing students' ability of clinical reasoning competency remains doubtful.

Under the current situations in Hong Kong, assessing nursing students' clinical reasoning competency has always been a challenge towards the clinical nursing educators; thus, there is a pressing need to develop a clinical assessment tool that can truly evaluate the clinical reasoning competency of the nursing students especially those who are in transition to qualified nurses.

1.3 Aim of the Study

The aim of this study is to establish a reliable and valid clinical assessment tool, named "Clinical Reasoning Competency Assessment Tool (CRCAT)", to examine the clinical reasoning competency of the Enrolled Nursing students.

1.4 Objectives of the Study

The objectives of this study are as follows:

i. To estimate the reliability of the clinical reasoning competency assessment tool (CRCAT) by Cronbach alpha statistics.

ii. To examine the construct validity of the clinical reasoning competency assessment tool (CRCAT).

iii. To examine the model-data fit of the clinical reasoning competency assessment tool (CRCAT).

iv. To evaluate clinical reasoning competency of participants.

1.5 Research questions

Reflected from background and the development of assessment tool in psychometric measure, this study asked the research questions as below:

1) What are the reliability and validity of the clinical reasoning competency assessment tool (CRCAT)?

2) What is the model-data fit of the clinical reasoning competency assessment tool (CRCAT)?

3) Are there significant differences in clinical reasoning competency between the final year Enrolled Nursing students and the clinical nursing instructors, and between the final year Enrolled Nursing students and the first year Enrolled Nursing students?

1.6 Research Hypothesis

1) There is a significant difference in clinical reasoning competency between the final year Enrolled Nursing students and the first year Enrolled Nursing students.

2) There is a significant difference in clinical reasoning competency between the clinical nursing instructors and the Enrolled Nursing students.

1.7 Significance of the Study

The study provides a model-data fit clinical reasoning assessment tool for assessing the clinical reasoning competency of Enrolled Nursing students, which examines not only the student ability but also the question item difficulty as well.

1.8 Definitions

The following definitions are extracted from literature:

1) Competency

Competency was referred to as an important capability that was needed to do a job; whereas, competence was used to describe the ability to do something well (Moghabghab, Tong, Hallaran and Anderson, 2018).

2) Clinical reasoning

Clinical reasoning was a professional and clinical context-dependent way of cognitive thinking process with decision making to guide clinical practice (Higgs, Loftus and Christensen, 2008).

1.9 Organization of the thesis

The chapters of this thesis are organized according to the stages of the study. After this introduction chapter, Chapter Two presents the literature review. A clear picture of how the assessment tool to be constructed, and what theory base used to categorize the level of clinical reasoning competency, the four building blocks of the conceptual framework for the development of the assessment tool as well as the partial credit model used as a tool of data analysis in this study are described in Chapter Three. The items, which are generated from the study for the construction as well as the establishment of content validity and understandability of the Clinical Reasoning Competence Assessment Tool (CRCAT) are illustrated and discussed in Chapter Four. The reliability and validity tests on the developed CRCAT are reported in Chapter Five. Chapter Six reviews, examines and discusses the usability of the assessment tool in the clinical area, with implications and limitations for future research. Finally, Chapter Seven summarizes the thesis and its main contributions.

1.10 Conclusion

In summary, this chapter has elaborated its background of the study on the development of clinical reasoning competency assessment tool (CRCAT), introduction, justification of the study, aim and objectives, research questions, research hypothesis, significance of the study, terms definitions, and organization of the thesis. The following Chapter Two presents the literature review.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

The literature review provided information over seven sections that corresponded to the concepts and frameworks presented within this thesis. The areas including competence, clinical reasoning, scripts, Bloom's educational objectives taxonomy and instrument development had been explored in the electronic databases of nursing, education, healthcare and medicine. Competence, clinical reasoning, competence in nursing, clinical reasoning in nursing, clinical reasoning assessment, competence assessment are key words and phrases used in the database searches. Ten databases had been searched including EBSCO, CINAHL, ERIC, PsycINFO, PubMed, Education Full Text (H.W.Wilson), Elsevier Sciencedirect-online Journals, Journals@Ovid and google scholar. The primary focus of the search was to locate articles regarding clinical reasoning in nursing and the assessment of clinical reasoning in nursing. Therefore, literature focusing on the teaching curricular in junior schools, high schools and schools for early childhood, strategies to enhance teaching and learning, and education assessments not for the healthcare profession were excluded. Inclusion criteria include literature focusing on competence and clinical reasoning in healthcare professionals, educational assessment, clinical competence assessments for healthcare professionals, Bloom's educational taxonomy and instruments development. These criteria were used to identify the appropriate methods to develop a valid and reliable assessment tool to evaluate the clinical reasoning competency of the nursing students as well as to examine the difference in competency between the groups of nursing students with different levels of study. The abstracts of total 1539 articles were reviewed; 319 articles were retrieved and reviewed; 72 articles met the inclusion criteria. In this chapter, related studies including the clinical reasoning and competence in nursing, script, problem solving inventory, Bloom's educational objectives taxonomy as well as validation of assessment tools had been described.

2.2 Clinical Reasoning

2.2.1. Definition of reasoning

Reasoning was studied from various points of views in different human sciences, such as Psychology, Sociology and behavioral sciences, reasoning was then defined based on the views of these different studies. Reasoning was a process of inferring solutions from a set of statements or related information, it was the process of applying logic thought patterns in the solution of problems (Walton,1990). Based on the studies of logic and psychology, reasoning could be defined as a process of inference from certain already known assumptions to other assumptions that were true but following from the previous assumptions, and it was the actual thought process of exercising the mind (Walton, 1990). The study of education also defined that reasoning was a process consisting of seven steps: observations, facts, inference, assumptions, opinions, arguments and analysis. In fact, people were always making observations, facts or theory was then established after observations. Inferences were drawn from the facts or theory; and several assumptions were then made from inferences. People used observations, facts or theory, inferences and assumptions to form opinions, and arguments created to defend many opinions. Finally, people would use analysis to critique their own observations, facts or theory, inferences, assumptions, opinions, arguments. Since reasoning was studied from various points of views in different human sciences, reasoning could be categorized into four types including non-dialectical and dialectical reasoning, alethic and epistemic reasoning, static and dynamic reasoning, practical and theoretical reasoning. Practical reasoning was a process of choosing a course of action based on goals, knowledge and uncertainty or a changing situation (Walton, 1990). In fact, Clinical reasoning was actually a kind of practical reasoning, and it is the main construct of the present study and would be further explored in the following section.

2.2.2. Definition of clinical reasoning

As explored in the above section, clinical reasoning was a kind of practical reasoning that was a process of selecting different appropriate actions based on the set goals and the known knowledge in a context with uncertainty. In the literature, multiple concepts had been used synonymously, such as decision-making, problem-solving, clinical judgment, diagnostic reasoning and clinical reasoning. While these concepts all included elements of both process and outcome, the concepts diagnostic reasoning and clinical reasoning focused on the thinking strategies that a nurse used to make a judgment or decision and solve problems (Murphy 2004, Kautz et al. 2005, Su et al. 2005, Simmons, B. 2010). There were many definitions of clinical reasoning given by different nurse leaders that were helpful to find the right research directions. Simmons (2010) defined clinical reasoning as a complex process that used cognition, metacognition, and discipline-specific knowledge to gather and analyze patient information, evaluate its significance, and weigh alternative actions. According to Victor-Chmil (2013), the terms critical thinking, clinical reasoning, and clinical judgment are interrelated concepts. Each of them had an important set of processes leading the nurses to implementing evidence-based practice. Critical thinking was the cognitive process used for analyzing knowledge, clinical reasoning was the cognitive and metacognitive processes used for analyzing knowledge related to a clinical situation or specific patient, and clinical judgment was the cognitive, psychomotor, and affective processes demonstrated through action and behaviors.

2.2.3. Clinical reasoning in nursing

Nurses with effective clinical reasoning skills had positive impacts on patient outcomes as clinical reasoning was one of the commonly used applications of critical thinking into the clinical situations, and it played a very important role and had a close relationship with nursing in the clinical environment. Clinical reasoning as a cognitive process could be used by healthcare practitioners to address patient health issues. Also, clinical reasoning involved synthesis of knowledge and experience requiring both a background of scientific knowledge and a clinical situation for application (Benner, 1984, Jones, 1988 & Lapkin et al., 2010, Victor-Chmil, 2013). Clinical reasoning could be defined as "the cognitive processes and strategies that nurses used to understand the significance of patient data, to identify and diagnose actual or potential patient problems, to make clinical decisions to assist in problem resolution, and to achieve positive patient outcomes" (Fonteyn & Ritter, 2008).

In fact, there were several models used to develop clinical reasoning ability for nursing students or nurses presented in research papers. Kautz, Kuiper, Pesut, Knight-Brown & Daneker (2005) used the Outcome Present State Test, OPT (Figure 2-1) and Self-Regulated Learning, SRL (Figure 2-2) strategies in their study to evaluate the clinical reasoning skills of the students through the application of OPT and SRL.

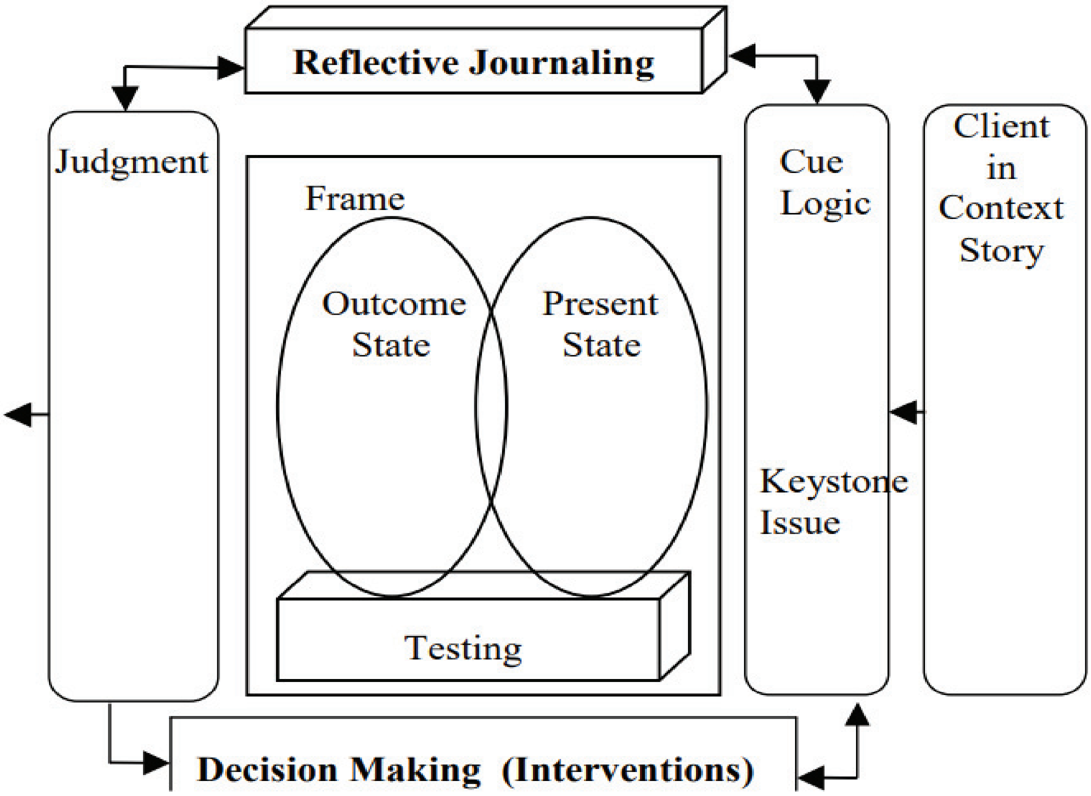

Figure 2-1: Outcome Present State Test (OPT) Model (Pesut & Herman 1999, Kautz, etc. 2005)

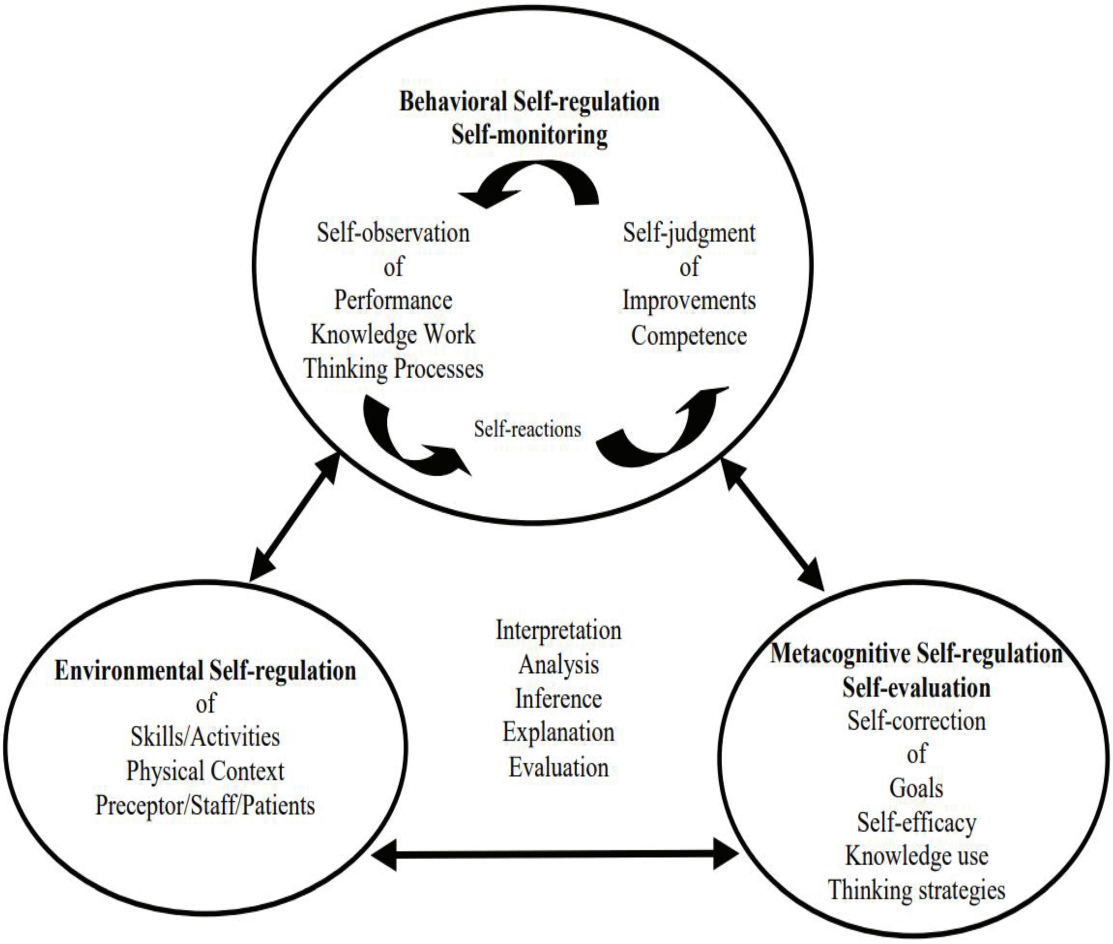

Figure 2-2: Reflective Self-regulated Learning in Nursing (Kuiper 1999, Kautz, etc. 2005)

The results revealed that students effectively made gains in learning associated with the OPT model, and they also made significant gains in self-observation, self-judgment, knowledge work and use of health care personnel resources during clinical experiences through the self-regulated learning prompt reflective journals. Hoffman et al (2010) used an Interactive Computerized Decision support Framework (ICDSF) that was designed based on the clinical reasoning cycle (Figure 2-3) consisting of the steps: consider, collect, process, identify, establish, act, evaluate and reflect to improve nurse students' specific knowledge and clinical reasoning skills. The results told that The ICDSF was useful in developing cognitive skills including clinical reasoning, problem-solving and decision-making. Levett-Jones et al (2010) developed a 'five rights educational model consisting of the right cues and to take the right action for the right patient at the right time and for the right reason that has the potential to improve nurse students' clinical reasoning skills and to increase their preparedness for professional practice. A framework for understanding clinical reasoning was developed and used by Carr (2004) in her study using interpretive research approach with multiple methods including focus groups, observation and narrative recordings for data collection and analysis. The four-stage framework including naming, framing, need identification and action options provided a template as an educational tool for the development of nurse students' clinical reasoning skills. Kuiper and Pesut (2004) revealed that the use of self-regulated learning strategies had supported and facilitated the learning of reflective clinical reasoning in nursing practice contexts. A model of clinical reasoning cycle (Figure 2-3) was also demonstrated in the Instructor Resources prepared by the faculty of Health at the University of Newcastle that consisting of eight main steps: considering, collecting, processing, identifying, establishing, acting, evaluating and reflecting, which is a dynamic process and can be combined one or more steps or move back and forth between them reaching the expected outcomes (Levett-Jones et al, 2010).

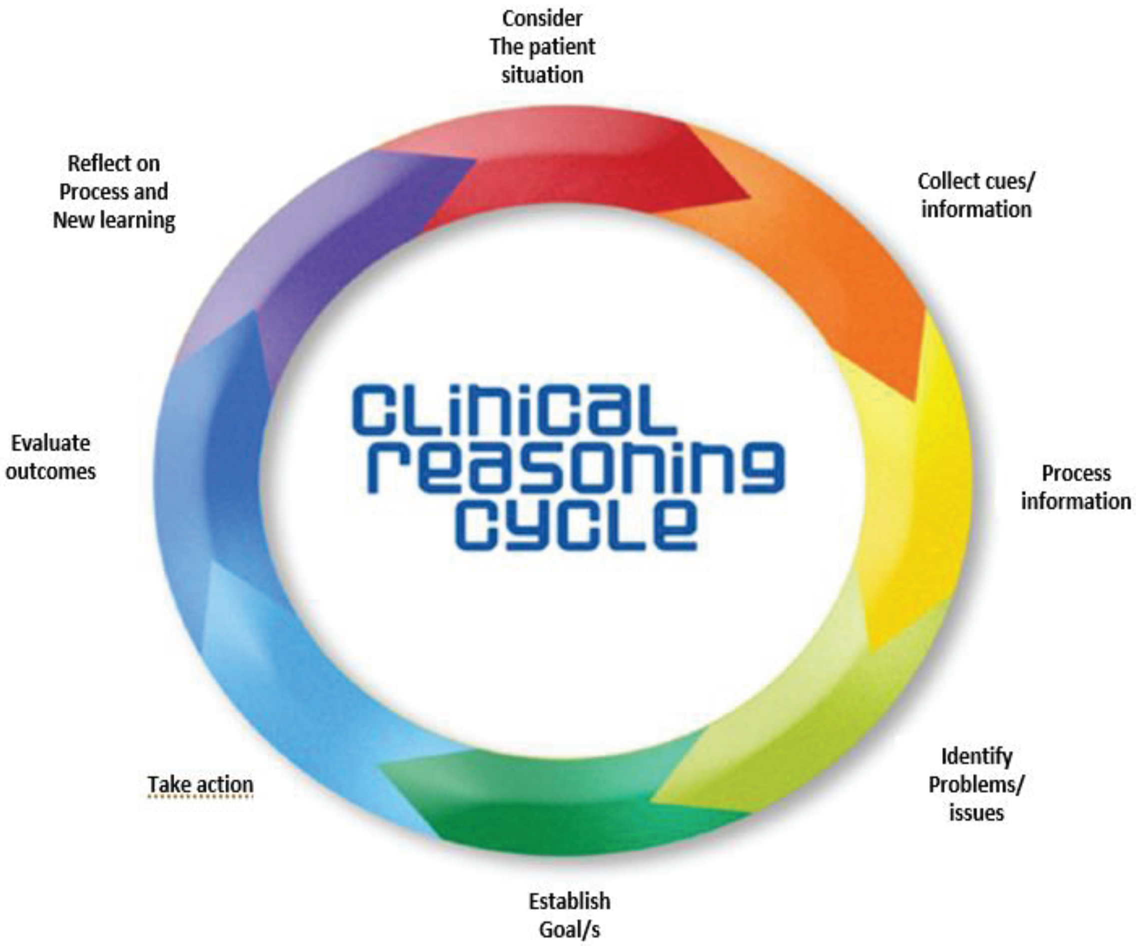

Figure 2-3: The clinical reasoning cycle (Levett-Jones et al, 2010)

This kind of clinical reasoning model may provide a learning environment guiding and monitoring the clinical reasoning process towards the expected positive clinical outcomes. In fact, the above said models have already been used as the teaching strategies in nursing education in order to develop and enhance nursing students' learning in the area of clinical reasoning during their training years. However, there is no assessment tool developed for evaluating the clinical reasoning ability of nurses. Since the clinical reasoning ability is a crucial clinical competence that nurses should always maintain in order to ensure safe practice, the aim of this study is to develop a clinical reasoning competence assessment tool that can be used to measure the competence level of the nurses in this area. The roles of clinical reasoning and competence in nursing are further explored in the following sections.

2.3 Clinical Competence

2.3.1. Definition of competence

Competence was the ability to consistently produce the work results that were required for the most efficient and effective achievement of the organizational goals or the professional standard (Teodorescu, 2006). Competence was the ability to use specified standards of skills and to apply relevant knowledge and understanding to the performance of relevant tasks and ultimately reached the set organizational or professional goals. In practice, the competence approach was applied by an individual in given context using skills, knowledge, and performances to carry out the required task successfully and safely. These were the elements or dimensions of competence (Ashworth & Saxton, 1990). The professional competence had several dimensions of competences, such as pedagogical competence includes teacher's competence in regard to collaboration, comprehensive view and contribution to the development of pedagogy for higher education. The concept of professional competence was used with the meaning of required professional standard that was often specified by law that raised a person in fulfilling the particular role of the profession (Ryegard, 2010, Suciu & Mata, 2011). In fact, according to the Nursing Council of Hong Kong, there were several core competence areas that the pre-enrolment nursing students were required to reach for the licensed recognition such as professional and legal ethical nursing practice, provision of care, personal and professional attribute, and teamwork.

2.3.2. Competence in nursing

Several systematic review studies found that in the last two decades of the 20^{th} Century, there were many controversies in clinical competence in nursing, the definition of competence had confused relationships with the terms of capability, expertise and performance. Lack of consensus was found among nursing professionals in that time, such as Benner (1982) defined that nursing competence was the ability to perform a task with desirable outcomes under different clinical situations in the real world. It was in the middle of a continuum ranging from novice to expert; whereas, Girot (1993) concluded that the definitions of nursing competence were divided into two main streams: one was behavioral-focused that was the ability to perform tasks, and the other one was psychological-focused that was including cognitive, affective and psychomotor skills. Although the definition of competence vary based on the level and experience of the health care providers, it was defined as a collection of knowledge, skills, and attitudes affecting and correlating with performance, and it could be measured using well-established standard (Klein, 2006). Since the growing of ethnic and racial diversity in the USA in the early 20th Century, culturally competence became a common but special concern in the nursing profession in the USA in order to address the issues that contributed to disparities in health care services. The cultural competence was described as the ability of health systems to provide care to people with diverse values, beliefs and behaviors (Brach, Hall & Fitall, 2019). For that, the primary goals of both the nurses' employers and the nurse trainers should be to deliver as well as to train high-quality and equitable health care providers to people, regardless of cultural background. In fact, up to early this Century, the definition of clinical competence in nursing has not yet come to consensus universally, furthermore, most of the methods in use to assess clinical competence in nursing has not been well developed, and without the tests of validity and reliability. Although the assessment of clinical competence in nursing remains the focus to pursue in the nursing education, there was still a gap between ideal and reality, and still no consensus on the definition of clinical competence in nursing profession. (Eraut, 1994 & 1998; Watson, Stimpson, Topping & Porock, 2002). In the first decade of this Century, the nursing professionals started to be aware of the critical role of competence-based approach in nursing education following the transition of nurse training form hospital-based to university-based training; however, there was little progress on consensus on the definition of competence in nursing practice. The definition of competence in nursing practice should be deriving from the holistic conception with the combination of knowledge, performance, skills, values and attitudes that can facilitate the development of competence standards as well as that of the assessment tools required for the clinical competence in nursing(Cowan, Norman & Coopama, 2005). In the second decade of this Century, competence became a word commonly used in the nursing profession universally though the concept was still not very clearly defined. Competence in nursing practice seemed to be the meaning of being slightly better than the time being newly qualified (Garside & Nhemachena, 2012). Watson et al. (2002) defined that competence was often no more than not being incompetence. Since the definition of competence remains nebulous and undefined, very few assessment tools that used to measure the competence in nursing have been developed. Nevertheless, a nursing faculty at a mid-western college adopted a competence assessment model developed by Lenbury, C. in 1999 as a guideline to develop clinical nursing competence assessment for each nursing procedures using corresponding critical elements (Klein, 2006) that seems to be similar to the clinical assessment criteria used in some schools of general nursing in Hong Kong nowadays; however, the competence assessment was procedure-oriented. As one of the health professionals, clinical competence of nurses should be maintained through continuing reflective practice, lifelong learning, and integration of learning into nursing practice (Bassendowski & Petrucka, 2009). Thus, clinical competence assessment needed to be implemented regularly for the nurses at different levels so as to ensure efficient and safe clinical practice as well as the benefits of the community. In fact, competence was the abilities individuals possess that enabled nurses to perform their duties so as to meet the required professional standard; however, these abilities might improve or diminish over time (Beidler, 2001, Tabari-Khomeiran, Kiger, Parsa-Yekta, & Ahmadi, 2007). The Competence Outcome Performance Assessment (COPA) Model was a comprehensive core competencies method designed model used to promote competence-focused and practice-focused nursing education. In the model, all nursing skills and abilities were clustered under eight core practice competence categories including assessment and intervention, communication, critical thinking or clinical reasoning, human caring relationships, teaching, management, leadership, and knowledge integration skills. These core competencies were universal (Armstrong, Spencer & Lenburg, 2009). Competence in nursing is not focusing on one area but consisting of several core competencies. Clinical reasoning is one of the core practice competencies; however, there are few assessment tools for measuring clinical reasoning competence used in the nursing profession as that suggested in the present study.

2.4 Clinical reasoning competency in nursing

2.4.1 Definition of competency

The terms competency and competence were used in similar ways to describe the ability to do something successfully or effectively. However, in specific interpretations, competency was used to describe as the ability to perform a certain task (Moghabghab, Tong, Hallaran and Anderson, 2018), moreover, competency had also been described as an important capability that was needed to do a job; whereas, competence was used to describe the ability to do something well (Moghabghab et al., 2018).

2.4.2 Clinical reasoning competency

Competences were centered on the individuals and are independent of the task-specific context in which performance occurs. Competence level was not only of a person but also of a context. People did not have competences independent of a context (Fischer et al.,1993; Le Deist & Winterton, 2005). In other words, the nursing professional competence were, of course, centered on nurses, and were independent of the clinical-specific context in which performance occurs, clinical competence level was not only of a nurse but also of the clinical context. No clinical context, no assessment of clinical competence. Competence were multi-dimensional frameworks including cognitive, skills and behaviors that was the commonly adopted competence-based training approach in the countries of Western Europe (Le Deist & Winterton, 2005). Moreover, professional competence was the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice for the most benefit of the individual and the community being served. Competence included three main functions: 1. Cognitive function using knowledge to solve clinical problems; 2. Integrative function using biomedical, nursing and psychosocial data in clinical reasoning; 3. Relational function using communication skills effectively to maintain therapeutic trust relationship (Epstein & Hundert, 2002). Therefore, clinical reasoning was one of the dimensions of clinical competence in health profession, furthermore, competency and competence were interchangeable, the former was used to describe the ability to perform a certain task; and the latter was the general ability (Moghabghab et al., 2018), hence, in this study, the clinical reasoning competency was to describe the ability of using thinking and decision making to guide professional practice in clinical context (Higgs, Jones, Loftus & Christensen, 2008).

2.4.3 Clinical reasoning competency and script

Charlin, Boshuizen, Custers and Feltovich (2007) mentioned that "Script" was a cognitive sciences concept aiming at explaining how people understand events that happen in the real world. Script activation was automatic and used in a strategic way to confirm hypotheses. The activated scripts served to guide information selection, memorization and interpretation; thus, script-based clinical reasoning was very efficient. Script concordance test (SCT) is based on the principle that the multiple judgments made in the clinical reasoning processes can be probed and their concordance with those of a panel reference experts can be measured that providing an assessment tool for clinical reasoning (Fournier, Demeester and Charlin, 2008). A script concordance test with 51 items scenarios and a total of 158 questions had been implemented by Duggan and Charlin (2012) as a summative assessment for the 5th year medical students' clinical reasoning, the results told that the script concordance test might be a useful method to assess clinical reasoning of the medical students with mean score 63.6 (7.6) and Cronbach coefficient alpha of 0.62 ($\rho < 0.05$). Humbert, Johnson, Miech, Friedberg, Grackin and Seidman (2011) had implemented a script concordance test for two groups of pre-clinical medical students studying in the 2nd year and the 4th year, the results showed that the SCT could be used as an assessment tool of measuring problem-solving performance in competency evaluation with Cronbach coefficient alpha 0.73 and the one-way ANOVA ($F_{2,508}$ =120.4; ρ <0.0001). Nouh et al. (2012) had implemented a study using a script concordance test as a measure of clinical reasoning for a total of 202 general surgical residents from different seniority enrolled across nine Canadian universities aiming to determine if the SCT maintained its validity and reliability when administered on a national level, the results showed that the SCT was a reliable and valid assessment tool with Cronbach coefficient alpha > 0.8 and was able to differentiate junior from senior with scores progressively increased. Meterissian, Zabolotny, Gagnon, & Charlin (2007) had examined the variability within the reference panel for a script concordance test, the results showed that the higher variability gave higher effect size for discrimination that implied that the variability of answers within the reference panel was a key component of discriminatory power of the SCT. Gagnon, Charlin, Coletti, Sauve and Vleuten (2005) had implemented a study in a group of 80 residents who were tested on 73 items with Cronbach's alpha 0.76, a total of 38 members made up the pool of experienced practitioners, from which 1000 random panels of reference of increasing sizes (5, 10, 15, 20, 25 and 30) were generated in order to examine the effect size of reference panel for a script concordance test, the results showed that there was a large difference for reliability in different panel size: a panel size of 5 (0.62), a panel size of 10 (0.70), a panel size over 20 (0.74) and 0.76 for the panel size of 38. The mean correlation coefficient values were 0.90 with 5 panel members, 0.95 with 10 members and 0.98 with 20 members. The panel size over 10 was associated with acceptable reliability and good correlation between the samples and the panel members. Hence, the reference panel with 15 to 20 members was advised in order to maintain high reliability and good correlation. Deschenes, Charlin, Gagnon and Goudreau (2011) found that the script concordance test was also a valid, reliable and standardized tool for assessing reasoning based on nursing care of patients with Cronbach's alpha coefficient of 0.86. In recent years, many research papers suggested using the SCT as a strategy for investigating the process of clinical reasoning within the health profession (Charlin and Vleuten, 2004). In fact, the script concordance test had been used in the medical profession field both education and clinical as an assessment tool for evaluating the clinical reasoning ability of the medical students as well as the medical clinicians since the year of 2000; however, script concordance test in nursing profession was still in its infancy, there was few similar studies implemented in nursing profession and no such assessment tool developed for evaluating the clinical reasoning for either the nurses or the nursing students in Hong Kong. This study had been designed to develop a clinical reasoning competency assessment tool (CRCAT) using script-based scenarios as an assessment tool used to evaluate the clinical reasoning competency of the pre-registration nursing students.

2.4.4 Clinical Reasoning Competency Assessment Tool and Problem-Solving Inventory

The Problem-solving Inventory (PSI) had been developed by Heppner and Petersen in 1982, and has been used in more than 100 studies and been referred to as one of the most commonly used self-report inventories to examining person's problem solving ability (Nezu & Perri, 1989; Heppner & Baker, 1997). The PSI was used to assess a person's awareness and evaluation of her or his problem solving abilities so as to recognizing if this person was a problem solver. The PSI was a self-report measure assessing perceptions of problem-solving rather than actual problem-solving skills (Heppner & Baker, 1997). The PSI consisted of 35 item-instrument that measured a person's perception regarding his or her problem-solving abilities and problem-solving style in daily living of his or her life. It consisted of three factors yielding three separate subscales: Factor 1. Problem-Solving Confidence, it contained 11 items assessing self-perceived confidence, belief and self-assurance in effectively solving problems (e.g. " I am usually able to think up creative and effective alternatives to solve a problem."). Factor 2. Approach-Avoidance Style, it contained 16 items assessing whether a person tended to approach or avoid problems (e.g. "When a solution to a problem was unsuccessful, I do not examine why it didn't work."). Factor 3. Personal Control, it contained 5 items assessing elements of self-control on emotions and behaviors (e.g. "I make snap judgments and later regret them."). All items were scored on a six-point Likert scale ranging from 1 = Strongly

Agree to 6 = Strongly Disagree. These three factors were not only intercorrelated but they have also been proven to be distinct dimensions (Heppner & Baker, 1997). The dimensions underlying the applied problem solving process and the problem solving inventory had been examined and the results revealed that the constructs were internally consistent with satisfactory reliability ($r = 0.89$) and estimates of validity suggested that the instrument was measuring constructs that were related to the general perceptions of problem-solving skills as well as the personality variables, most notably personal control, and all correlations were statistically significant ($p = 0.01$). Since the developed clinical reasoning competency assessment tool (CRCAT) had been designed for measuring one construct of clinical reasoning competency that could be defined as "the cognitive processes and strategies that nurses use to understand the significance of patient data, to identify and diagnose actual or potential patient problems, to make clinical decisions to assist in problem resolution, and to achieve positive patient outcomes" (Fonteyn and Ritter, 2008). Therefore, both the PSI and the CRCAT were the tools being used to assess problem-solving related ability but the former one was perception-focused and the focus of the latter one was cognitive process. Since there was few similar clinical assessment test using scripts and Bloom's revised levels of cognitive as indicators of clinical reasoning competency had been developed to evaluating the clinical reasoning competency of nursing students, the problem solving inventory (PSI) and the developed clinical reasoning competency assessment tool (CRCAT) were then being used at the same time in the present study to examining the construct validity of the developed CRCAT.

2.5 Bloom's Educational Objectives Taxonomy

2.5.1. Bloom's Original Taxonomy

According to Munzenmaier and Rubin (2013), Benjamin Bloom has begun developing his taxonomy since 1948 starting from a series of informal discussion with his colleagues at the American Psychological Association in 1948 aiming at formulating assessments to measure learning using a group of common languages that could be used by the educational measurement experts especially for sharing findings and exchanging assessments items. Bloom had convinced his collaborators including educators, instructional designers, researchers, and evaluators to organize the educational goals and the learning behaviors of the taxonomy into a hierarchy from the simplest to the most complex on a continuum. For that, the taxonomy was named as Bloom's Taxonomy of Educational Objectives because of Bloom's foundational contribution to this project. The first edition was published in 1956. It consisted of three domains including cognitive domain with knowledge-based; affective domain with attitude-based and the psychomotor, which was physical skills-based domain. The cognitive domain had six levels organized in a hierarchy from the lowest level of knowledge going through the levels of comprehension, application, analysis, synthesis to the highest level of evaluation. The lowest level of knowledge was the foundation of all cognition. The level of cognition domain ascended implying the cognitive ability increased.

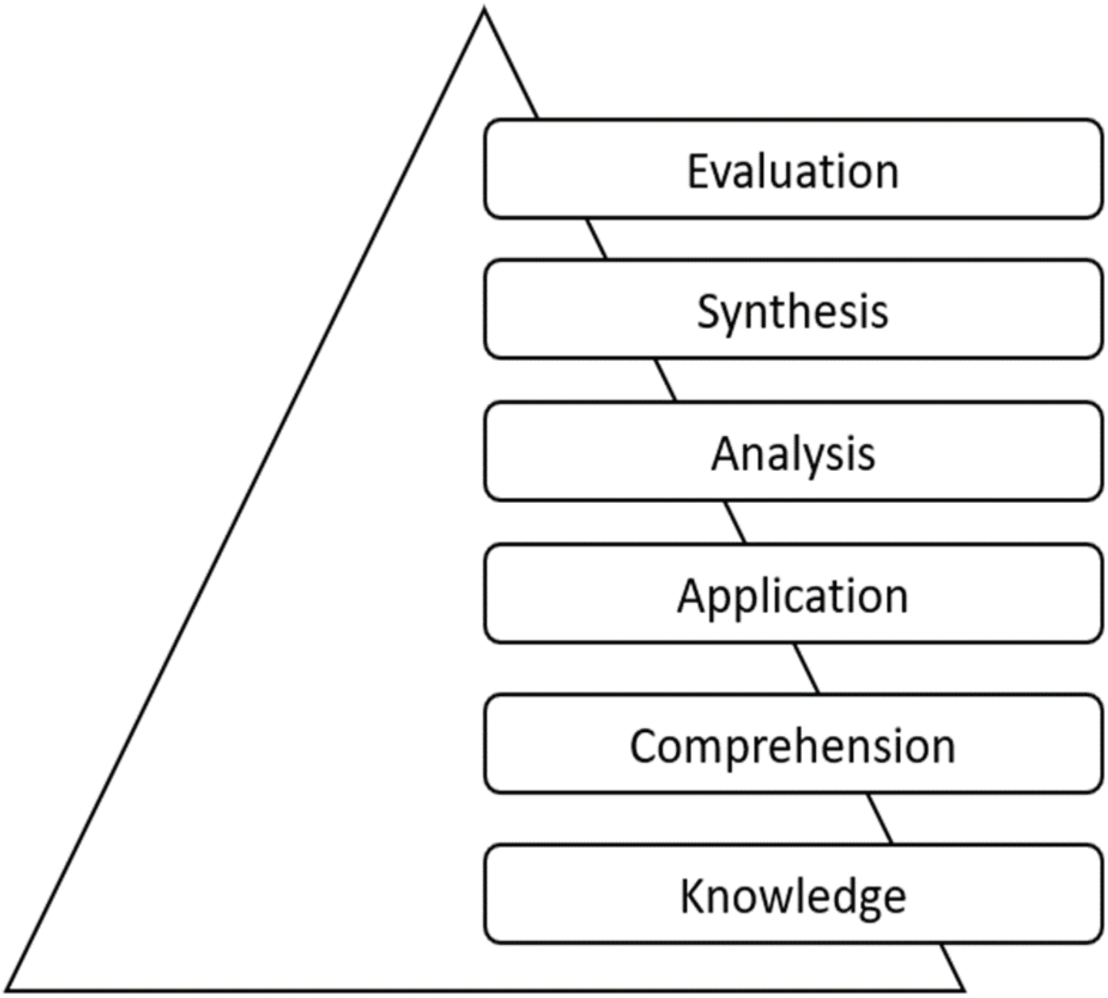

Figure 2-4: The Bloom's original taxonomy (Munzenmaier & Rubin, 2013)

The knowledge level that is at the bottom of the hierarchical triangle (Figure 2-4) was defined as remembering or retrieving previously learned information. The comprehension level that was the second level from the bottom was defined as understanding the meaning of the previously learned information. The application level that was third level from the bottom was defined as using the information in a situation. The analysis level was the fourth level from the bottom was defined as breaking information into parts to understand it more fully. The fifth level was the level just below the highest was defined as putting ideas together to form something new. The sixth level was the highest level that was defined as making judgments. Knowledge and comprehension were often referred to as lower-order thinking skills; whereas, the skills above were termed higher-order or critical thinking skills. According to the ideas of the hierarchy, if the learning objectives focused mainly on the first and second levels, learners might understand what they had learned but failed to recognize the appropriate time to apply their knowledge. Therefore, the higher-order learning objectives could facilitate the learners to develop the higher cognitive ability to identify critical variables and make appropriate judgments to solving problems. Nowadays, Bloom's taxonomy with six levels of different complexity is still the most widely used in the field of education profession.

2.5.2 Validity of Bloom's Original Taxonomy

A meta-analysis study had been done by Seddon (1978) aiming at examining the validity of Bloom's taxonomy; however, the results of the research papers that being analyzed were so varied that could not reach any recommendations on the validity of Bloom's taxonomy but concluded that no one had been able to demonstrate that the properties of Bloom's taxonomy did not exist; and also no one had demonstrate that they did (Seddon, 1978). In response to the study of Furst (1981), it advocated a continued need for the classification of cognitive objectives and acknowledged that there was no single scheme emerged as an all-inclusive or all-purpose tool. Since then, there was no more research and essays regarding the validity of Bloom's taxonomy appeared (Seaman, 2011).

2.5.3 Bloom's Revised Taxonomy

A revision of Bloom's taxonomy of educational objectives had been prepared and published by Anderson, L. and his collaborators in 2001 in order to meet the needs of the education profession (Munzenmaier & Rubin, 2013). In the revised taxonomy, the category names are no longer nouns but verbs. Evaluation replaces synthesis as the second high level of the pyramid; and a new category of creating was placed at the highest level (Figure 2-5).

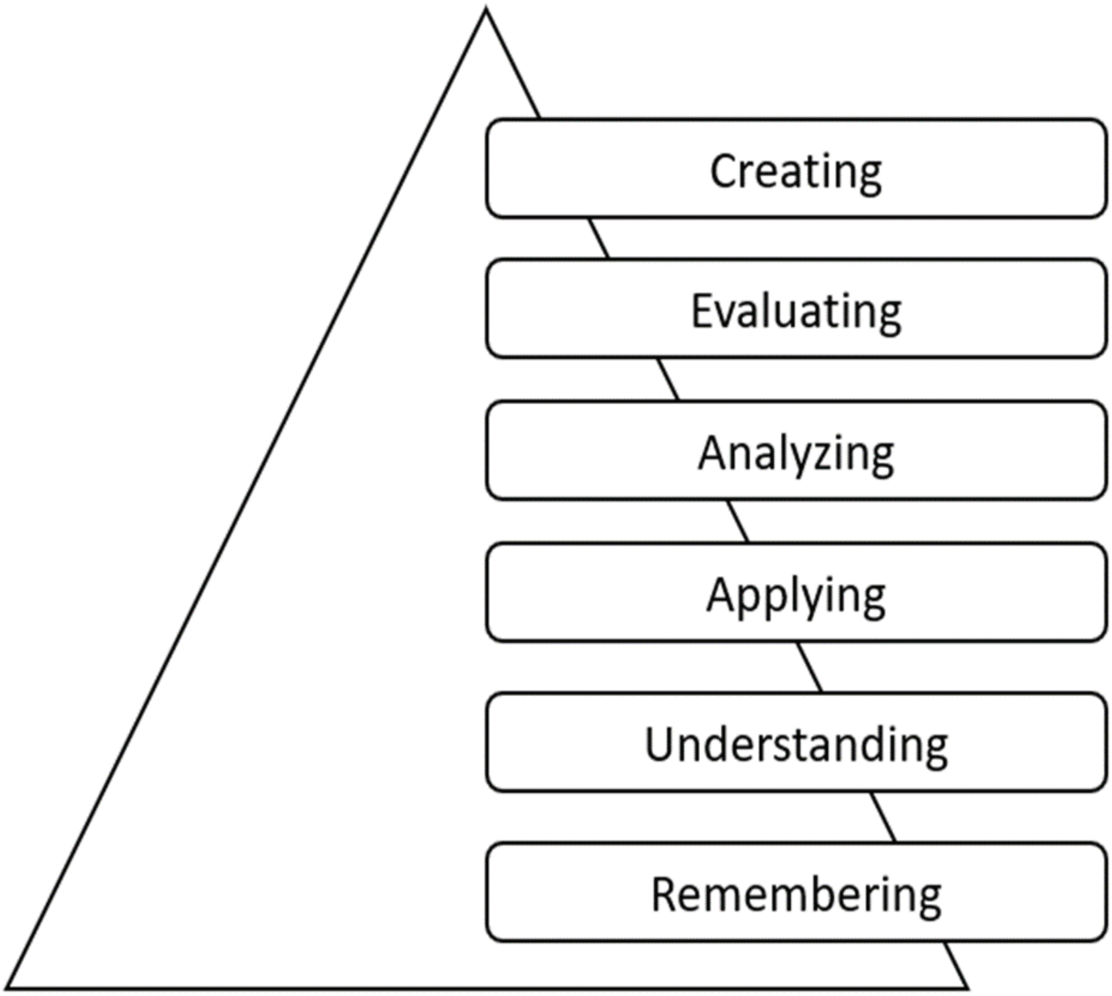

Figure 2-5: Bloom's revised taxonomies (Munzenmaier & Rubin, 2013)

In the original taxonomy, the six categories were arranged in a hierarchical pyramid, and the learners were required to master the lowest level of the hierarchy before they could go to the next higher level; whereas, the revised taxonomy also has similar arrangement but overlapping of categories was allowed, which is the cognitive processes organized along a continuum from the most basic to the most complex (Table 2-1) that was one of the dimensions of the revised taxonomy. The revised Bloom's taxonomy based on the structure of the original one had been reorganized to form two dimensions of knowledge and cognitive process. The dimension of cognitive process with six levels of complexity was retained in the revised version, the verb aspect of original Knowledge was renamed "Remember", Comprehension was renamed "Understand". Application, Analysis, and Evaluation were retained but in verb forms as Apply, Analyze, and Evaluate. Synthesis changed places with Evaluation that was renamed Create (Anderson & Krathwohl, 2001).

Table 2-1: Revised Bloom's Taxonomy Action Verbs

Definition	I. Remembering	II. Understanding	III. Applying	IV. Analyzing	V. Evaluating	VI. Creating
Bloom's definition	Exhibit memory of previously learned material by recalling facts, terms, basic concepts, and answers.	Demonstrate understanding of facts and ideas by organizing, comparing, translating, interpreting, giving descriptions, and stating main ideas.	Solve problems to new situations by applying acquired knowledge, facts, techniques and rules in a different way.	Examine and break information into parts by identifying motives or causes. Make inferences and find evidence to support generalizations.	Present and defend opinions by making judgments about information, validity of ideas, or quality of work based on a set of criteria.	Compile information together in a different way by combining elements in a new pattern or proposing alternative solutions.
Verbs	Choose Define Find How Label List Match Name Omit Recall Relate Select Show Spell Tell What When Where Which Who Why	Classify Compare Contrast Demonstrate Explain Extend Illustrate Infer Interpret Outline Relate Rephrase Show Summarize Translate	Apply Build Choose Construct Develop Experiment with Identify Interview Make use of Model Organize Plan Select Solve Utilize	Analyse Assume Categorize Classify Compare Conclusion Contrast Discover Dissect Distinguish Divide Examine Function Inference Inspect List Motive Relationships Simplify Survey Take part in Test for Theme	Agree Appraise Assess Award Choose Compare Conclude Criteria Criticize Decide Deduct Defend Determine Disprove Estimate Evaluate Explain Importance Influence Interpret Judge Justify Mark Measure Opinion Perceive Prioritize Prove Rate Recommend Rule on Select Support Value	Adapt Build Change Choose Combine Compile Compose Construct Create Delete Design Develop Discuss Elaborate Estimate Formulate Happen Imagine Improve Invent Make up Maximize Minimize Modify Original Originate Plan Predict Propose Solution Solve Suppose Test Theory

The other dimension is at the knowledge level that contains four subcategories arranged from the most concrete to the most abstract including factual, conceptual, procedural and metacognitive (Table 2-2).

Table 2-2: Revised Bloom's Knowledge Dimension

Dimension	Definition
Factual Knowledge	The basic elements students must know to be acquainted with a discipline or solve problems in it.
Conceptual Knowledge	The interrelationships among the basic elements within a larger structure that enable them to function together.
Procedural Knowledge	How to do something, methods of inquiry, and criteria for using skills, algorithms, techniques, and methods.
Metacognitive Knowledge	Knowledge of cognition in general as well as awareness and knowledge of one's own cognition.

2.5.4. Bloom's Original Taxonomy and the Revised Bloom's Taxonomy

Since the revised Bloom's taxonomy was a two dimensional framework including knowledge and cognitive processes that could facilitate the nurse educators to develop lesson plans, to plan a program or entire curriculum as well as to examine the existing program or curricula; also, it could help the nurse educators making informed decisions about where and how to improve instructional design and delivery if compared with the Bloom's original taxonomy (Su & Osisek, 2011). Thus, the revised Bloom's taxonomy was used in the present study.

2.5.5. Bloom's Educational Objectives Taxonomy and Competence

Competence could be defined as the capacity to identify a problem and act skillfully in solving the problem. In dental education, the area of competencies had included the three domains of Bloom's original taxonomy, which were cognitive, affective and psychomotor (Beltran-Aguilar & Beltran-Neira, 2004). In nursing education, the three domains of Bloom's taxonomy were commonly used to identify and evaluate student clinical performance. Since Bloom's cognitive skills for educational measurement constitute six category levels of mental abilities that were from simple to complex, the cognitive domain was then predominantly used to write clinical learning objectives for evaluating student clinical ability (Field, Gallman, Nicholson, & Dreher, 1984). In fact, either the original version or the revised version of Bloom's taxonomy, the six levels of cognitive process were actually involved in clinical reasoning that was the pivotal cognitive activity of a practicing health professional making a clinical diagnosis that led to appropriate and safe clinical intervention (Nkanginieme, 1997). Therefore, the cognitive process from the first level to the fifth level including remembering, understanding, applying, analyzing, evaluating in the revised Bloom's taxonomy was used as a continuum competency level of clinical reasoning for the development of the clinical reasoning competency assessment tool (CRCAT) in the present study.

2.6 Validation of Assessment Tools

Cerit, Keskin and Ekici (2018) had implemented their study of development a valid, reliable, short and comprehensive bullying behaviours in nursing education scale to measure the bullying behaviours of nursing students in the education environment using a randomized with volunteer-based. A two-part form was used to collect data. The first part was to collect the demographic characteristics of nursing students. The second part was the development of bullying behaviours in nursing education scale. LISREL 8.80 and SPSS 20.0 software were used for data analysis. The Cronbach alpha coefficient was 0.88 and structure reliability was 0.92 showing that the bullying behaviours in nursing education scale was reliable and valid; however, data had been collected from the nursing students that was limited to the nursing students only, also, participation was low with 23% of nursing students from the two universities joined the study. Thus, further study with larger sample size and extended scope of sample was necessary in order to ensure the representative and generalization of the study results. Sun, Arning, Bochmann, Borger & Heitmann (2018) has developed and evaluated the reliability of the occupational safety and health monitoring and assessment tool (OSH-MAT) in cross-sectional survey among 128 companies as well as its validity among 30514 companies. Inter-rater reliability was examined in a cross-sectional survey among 128 companies and analyzed using interclass correlation coefficient (ICC). Content and construct validity were evaluated using the routine documented OSH-MAT values and injury rates at 30514 companies. Poisson regression analysis used to examine the content validity. Construct validity was examined using the principal component factor analysis with varimax rotation. Results of ICC was between 0.64 and 0.74 indicated good to very good inter-rater reliability of OSH-MAT values. Factor analysis identified three component subscales that met exactly the structure theory of this instrument. The Poisson regression analysis showed statistically significant exposure response relationship between OSH-MAT values and the five years average injury rates. The results showed that OSH-MAT was a valid and reliable instrument that could be used; however, this OSH-MAT instrument was used for injury prevention but not for early recognition and elimination of all risks to workers' lives and health at work, so, further study was necessary, such as at which level of OSH-MAT was an intervention program more effective, and to which extent of changing OSH-MAT values might largely improve the injury rates still needed to be investigated in future study. Siegle and Cardoso de Sa (2018) has used cross-sectional, prospective and descriptive approach in their study aiming at verifying the concurrent validity of Alberta infant Motor Scale in infants exposed to HIV, and the correlation of AIMS and Bayley Scale for this population as well as to comparing if these coefficient differed in the central age groups and extremities of AIMS. 82 infants exposed to HIV evaluated in different months of life (1st, 2nd 3rd 4th 8th 12th 15th 16th 17th 18th), with Alberta infant motor scale and Bayley scale. Results of raw cores of scales were compared with the correlation analysis. First concurrent analysis was Alberta score with Bayley's total motor score, and second analysis was Alberta score with Bayley's gross motor score. Pearson coefficient analysis was done and results showing that correlation with higher coefficient value only between the gross motor skills. Correlations were lower up to four months of age that was the limitation that it was because of the scale characteristics and the lower motor repertoire up to four months of age. Guine et al. (2016) has implemented a cross-sectional study to develop and validate an instrument to evaluate the knowledge of the general population about dietary fibers. A developed questionnaire of self-response was used to collect the data from the participants who were residing in 10 countries from three continents aiming at assessing their knowledge about dietary fibers. Exploratory factor analysis was chosen as the analysis of the main components using varimax orthogonal rotation and eigenvalues greater than 1.0. In the confirmatory factor analysis by structural equation modelling was considered the covariance matrix and adopted maximum likelihood estimation algorithm for parameter estimation. The exploratory factor analysis retained two factors. The first one called the dietary fibers and promotion of health that included seven questions with alpha 0.852. The second one was the sources of dietary fibers that had four questions with alpha 0.786. The results provided good internal consistency by the values of composite reliability 0.854 and 0.787 showing that the knowledge of dietary fibers scale was reliable and valid. Since the participants were residing in 10 different countries, the use of different languages could be a source of variability, and the diverse cultural environment could also create some variability that might affect the results. Farra, Smith, French & Gillespie (2015) has developed an assessment instrument to evaluate performance of the nursing skills of decontamination using a cross-sectional approach. All participants who were the students studying in two colleges of nursing were evaluated by a rater trained by one of the researchers to use the decontamination rubric. The performance of the participants was rated using the developed instrument. Forty-one participants were evaluated by two raters in order to assess the inter-rater reliability. The content validity index for the overall instrument score was 0.94. Internal consistency coefficient was 0.607. The inter-rater reliability was 0.9114. The results showed that the developed instrument was reliable and valid. But the sample size was small comparatively and the scope was narrow, also, the instrument was assessed in simulated settings and not in an actual disaster environment that could not reflect the actual performance of the nursing students or the nursing staff during the real or even disaster conditions. Waehrens and Fisher (2009) has implemented a pilot study aiming to explore the possibility of developing an instrument of linear measures of ADL ability based on the ADL taxonomy and a three-category rating scale. Data were obtained from medical records of the participants with moderate to severe brain injury and rated on the ADL taxonomy and recorded on the ADL taxonomy circle based on direct observations. Rasch rating scale model was applied to examining the possibility of converting the raw ordinal scores into equal-interval estimates of the ADL ability of the participants. The rating scale used in this pilot study displayed sound psychometric properties. Items from the ADL defined one construct. The item difficulty hierarchy was the same as in the hierarchy originally published that was supporting the reliability of the item difficulty estimates. The ADL taxonomy discriminated well between the different levels of ability in the sample of participants with brain injury however, the results were only based on participants with moderate to severe brain injury, they were rated on items relevance to their everyday lives only, also, the number of observations on difficult items was limited; so, replication of the study findings on a larger and more diverse scope of sample was necessary. Jones et al. (2009) had developed and validated a COPD assessment test using a cross-sectional approach. Twenty-one candidate items that were identified through a qualitative research with COPD patients were used in three prospective international studies. Psychometric and Rasch analysis identified eight items fitting the unidimensional model to form the CAT, items were tested for differential functioning between countries using Cronbach's alpha coefficient and ICC. Results showed that the COPD assessment test (CAT) had good measurement properties with Cronbach alpha coefficient at 0.88 and ICC at 0.8. However, it could not provide a valid, reliable and standardized measurement assessment of COPD health status with worldwide relevance since the reliability and validity findings were based on the data collected from the USA only. Moreover, the CAT should improve communication between clinician and participants in order to enable a treatment better targeted and managed. Andrew et al. (2007) has developed and validated the children asthma control test (C-ACT) using cross- sectional approach. A questionnaire with twenty-one items had been administered to 343 children aged 4 to 11 years old with asthma and their caregivers randomly assigned (75%, n=257) for development and cross-sectional validation of the tool and (25%, n=86) to a confirmatory sample. Stepwise logistic regression was used to reduce twenty-one items to those best able to discriminate control as defined by the specialist rating of asthma control. Seven items were selected from regression analysis of the development sample to comprise the children asthma control test. Scores of each item were summed for a total score with lower scores indicating poorer control. Summed scores discriminated between groups of participants differing in the specialist rating asthma control ($F=36.89$; $p<0.0001$), the need for changing in participants' therapy ($F=20.07$; $p<0.0001$). A score of 19 indicated inadequately controlled asthma (specificity 74%, sensitivity 68%). These analyses were confirmed in the confirmatory sample. Thus, the C-ACT was a validated tool and was able to discriminate among various levels of control. This study captured a relatively mild and controlled sample of children with asthma that might limit the generalizability of the study findings to other cohorts. Further study with a larger and diverse scope of samples was necessary. Fuchs-Lacelle and Hadjistavripoulos (2004) has developed a clinical assessment checklist for the seniors with limited ability to communicate aiming to assess pain in seniors with severe dementia. The study had three phases. In phase one, twenty-eight caregivers were interviewed in order to generate a list of pain-related behaviours for development of a pain assessment checklist for seniors with limited ability to communicate. Forty caregivers had completed the assessment checklist for the internal consistency test in phase two and thirty-four caregivers had completed the pain assessment checklist for the preliminary validation test in phase three. A pain assessment checklist for seniors with limited ability to communicate was created in phase one. In phase two, Cronbach's coefficient alpha was checked at 0.92 indicating that the pain assessment checklist had good internal consistency. In phase three, the pain assessment checklist was shown to be significant with ANOVA less than 0.001, and the concurrent validity was also tested for the correlations between the global intensity ratings and the pain assessment checklist for seniors with limited ability to communicate (PACSLAC) with significant less than 0.05. However, caregivers provided the retrospective reports about the events of painful, calm or distressing that would create probability of memory bias. The data collected were based on non-standardized naturally occurring situations that needed further assessment of the validity of the PACSLAC when the pain stimulus was kept consistent across the disabled seniors. Warden, Hurley & Volicer (2003) has developed a pain assessment in advanced dementia (PAINAD) scale using expert clinicians and behaviours observation methods. The study had two phases. A five-item observational tool with a range of 0 to 10 was developed in phase one and it was compared with the discomfort scale and two visual analog scales by trained raters and expert clinicians in phase two. Results showed that the Cronbach's alpha was lower than the desired 0.70 for the PAINAD scale but significant correlations were found with the discomfort scale-dementia of Alzheimer type (DS-DAT) that provided evidence of construct validity. Also, the PAINAD detected statistically significant differences between scores obtained before and after receiving pain medication. Results showed that the PAINAD was a simple, valid, and reliable instrument for measurement of pain in non-communicative participants. However, small subjects sample included only white, elderly, male at middle class, also, sample scope was narrow that the generalizability and representative of the results of this study was poor.

2.8 Conclusion

In summary, the findings obtained from the literature researching about the assessment tools for assessing the clinical competency in the nursing field substantiate the need for this study. For example, there were very few clinical assessment tools that were used to measure the clinical reasoning competency in nursing had been developed. Although a competency assessment model has been developed by Lenbury, C. in 1999 was used as a guideline to develop clinical nursing competence assessment but for the nursing procedures only (Klein, 2006). It seems similar to that has usually been used in the hospital-based schools of nursing in Hong Kong, which was actually procedure-oriented and raw score calculation. In fact, the clinical reasoning competency was actually the capability that each nurse should possess that enabling her or him to perform own duty up to the required professional standard; however, this kind of clinical capability may improve or diminish over time (Beidler, 2001; Tabari-khomeiran and Parsa-Yekta, 2007). The literature demonstrated a lack of consensus among the nursing professionals universally on the clinical competency in nursing. In fact, up to early this

Century, the universal definition of clinical competence in nursing had not yet come to a consensus. Although the focus of assessing clinical competence in nursing remained to pursue in nursing education, there is still a gap between ideal and reality, and still no consensus on the definition of clinical competence in the nursing profession. (Eraut, 1994 & 1998; Watson, Stimpson, Topping and Porock, 2002). Nevertheless, competence was centered on the individuals, and was independent of the task-specific context in which performance occurs. Competence level was not only of a person but also of a context. People did not have competences independent of a context (Fischer et al.,1993; Le Deist and Winterton, 2005). Both the original version and the revised version of Bloom's taxonomy, the six levels of cognitive process were actually involved in the process of clinical reasoning that was the pivotal cognitive activity of a practicing health professional making a clinical diagnosis that led to appropriate and safe clinical intervention (Nkanginieme, 1997). Therefore, the cognitive process from the first level to the fifth level including remembering, understanding, applying, analyzing, evaluating in the revised Bloom's taxonomy was used as a continuum competency level of clinical reasoning for the clinical reasoning competency assessment tool (CRCAT) in the study. Moreover, since there was less information regarding clinical reasoning competency tests in nursing and the problem solving inventory (PSI) was the tool used to assess problem-solving related cognitive ability, thus, in this study, the PSI and CRCAT were used at the same time in order to test the construct validity of the CRCAT.

The literature substantiated that all the newly developed measurement instruments or assessment tools should have been thoroughly examined for their reliability and validity before being used. This study has attempted to fill the gaps in the literature by reporting the validation of the developed clinical reasoning competency assessment tool (CRCAT) and its impacts on nursing profession including nursing administration, nursing research, nursing education and nursing practice.

CHAPTER THREE

STUDY METHODOLOGY

3.1 Introduction

This chapter described the methodology of study. This study was divided into two phases. Phase one was the development of clinical reasoning competency assessment tool (CRCAT) and phase two was the establishment of CRCAT through validation of the reliability and validity of CRCAT.

3.2 Phase one – Development of CRCAT

Phase one was the development of clinical reasoning competency assessment tool (CRCAT) using the conceptual framework of development model for measurement suggested by Wilson (2005), which consisted of four building blocks including construct map, item design, outcome space and measurement model. The content validity and understandability of the CRCAT were established in this phase. The details were described in the following Chapter Four.

3.3 Phase two – Establishment of CRCAT

The aim of this phase is to validate the clinical reasoning competency assessment tool (CRCAT). Through the implementation of the developed CRCAT to successfully achieve the following objectives:

i. To estimate the reliability of CRCAT by Cronbach alpha statistics.

ii. To examine the construct validity of CRCAT.

Iii To examine the model-data fit of CRCAT

iv. To evaluate clinical reasoning competency of participants

3.4 Research questions

The research questions are as follows:

1). What are the reliability and validity of the clinical reasoning competency assessment tool (CRCAT)?

2). What is the model-data fit of the clinical reasoning competency assessment tool (CRCAT)?

3). Are there significant differences in clinical reasoning competency between the final year Enrolled Nursing students and clinical nursing instructors, and between the final year Enrolled Nursing students and the first year Enrolled Nursing students?

3.5 Study Design

It was a cross-sectional design with exploration of data from the Enrolled Nursing students having different years of clinical experience to determine the relationships with their competency of clinical reasoning in clinical practice.

3.6 Method

3.6.1 Subjects

This study had two groups of Enrolled Nursing students and one group of clinical nursing instructors. The first group was the first year Enrolled Nursing students, and the second group was the final year Enrolled Nursing students who was in transition to Enrolled Nurses. They were drawn from a school of general nursing training in Hong Kong. The third group was the clinical nursing instructors being drawn from the medical wards and the surgical wards of the hospitals where the Enrolled Nursing students had their clinical practicum. All participants were adults living in Hong Kong for at least seven years. The clinical nursing instructors were registered nurses with more than 5 years of clinical experience working in both medical wards and surgical wards as advanced practicing nurses, they also played the role as clinical nursing instructors to coach the Enrolled Nursing students having practicum in hospitals.

3.6.2. Sampling

Since there is only one hospital-based school of general nursing under the Hong Kong Hospital Authority responsible for training Enrolled nurses (two-year Enrolled nurses training program), all the Enrolled Nursing students in this school who met the inclusion criteria were invited. 41 first year Enrolled Nursing students and 41 final year Enrolled Nursing students who were in transition to Enrolled nurses were willing to participate in this study. Also, the clinical nursing instructors who were advanced practicing nurses working either in medical wards or in surgical wards were invited to participate in this study. 82 Enrolled Nursing students and 15 clinical nursing instructors were recruited by a convenience sampling method.

3.6.3 Sample size

According to the Rasch measurement experts, a useful exploratory work using Rasch analysis with a small sample can certainly be performed. A sample size of 50 well-targeted examinees is conservative and acceptable for obtaining useful and stable estimates (Linacre, 1994). The sample size of 97 well-targeted participants was used in this study.

3.6.4. Instruments

The instruments used in this study were the developed clinical reasoning competence assessment tool (CRCAT) and the problem- solving inventory (PSI). The latter, with its established validity and reliability, was used as a standard for comparison.

3.7 Procedure

All sample subjects including the two groups of Enrolled Nursing students and the clinical nursing instructors signed the informed consents for themselves after full explanation on the aim, objectives, process and method of the study given. Each of them was provided an information sheet. They completed the CRCAT and PSI within the allowed time. For the two groups of Enrolled Nursing students, the study was conducted in the classroom; whereas, the clinical nursing instructors was at their office.

3.8 Ethical consideration

The rights, needs, values and desires of the participants were respected and safeguarded. The aim of the present study had been articulated verbally and in writing so that they were clearly understood by the participants. An explanation sheet was given to each participant and an informed consent was also collected from each of the participant before the study started. The participants were informed of all the data collection devices and activities. The participants were informed that all the collected personal data would be kept strictly confidential. The participants' rights, interests and wishes were considered first when choices were to be made regarding reporting the data. Code numbers were used instead of names to identify the questionnaires. The participants were informed that all the names and code numbers would be destroyed upon the completion of data collection. Written permissions were applied for the present study through the concerned research committees looking after the research activities.

Both the Research Ethical Approval Letter from the Education University of Hong Kong Human Research Ethics Committee (HREC) and the Research Approval Letter from the University of Hong Kong and Hong Kong West Cluster (research site) Clinical Research Ethics Review (Institutional Review Board of HKU / HAHKWC) were obtained.

3.9 Data Analysis - Validity and reliability

3.9.1 Validity

Measurement of validity was concerning the extent to which the new assessment tool measured what it was intended to measure. Also, the assessment tool was usually devised for purposes of discrimination, evaluation, or prediction. Validity implied that a measurement was relatively free from error, thus, a valid assessment tool was also reliable. Validity was to be evaluated within the context of the new assessment tool which was intended use. An assessment tool was used to make inferences about the magnitude of a particularly variable based on a relevant observable behavior or response, and validity was the basic to establishing these inferences. Validation was a process of hypothesis testing, determining if the test scores were related to specific behaviors, characteristics or levels of performance. Evidences to support hypothesis was generally defined according to four types of validity including face validity, content validity, criterion-related validity and construct validity (Portney & Watkins, 2009). Since the face validity was the weakest form of measurement validity, it was not used to test the validity of CRCAT. The content validity of CRCAT was established in the following Chapter Four.

Criterion-related validity is the most practical and objective approach to validity testing, and it has two components of concurrent validity and predictive validity. Concurrent validity was established when the criterion measures were taken at relatively the same time so that they both reflected the same incident of behaviors. Predictive validity was established when the outcome of the target test could be used to predict a future outcome. Construct validity was to test the ability of the target instrument to measure an abstract construct and the degree to which the instrument reflected the theoretical components of the construct (Portney & Watkins, 2009). Since there was no clinical assessment tool that was developed previously for assessing the clinical reasoning competency of nursing students, the concurrent validity could not be tested in this study. Nevertheless, Rasch measurement model analysis and groups comparison using ANOVA was used to examine the construct validity of the CRCAT instead.

3.9.2 Reliability

Reliability is the first prerequisite of a useful measurement tool. It was used to test the extent to which a measurement was consistent and free from error so that the data collected could be relied as accurate and meaningful indicators of an attribute to be assessed by the measurement (Portney & Watkins, 2009). Therefore, in this study, test re-test reliability and Rasch measurement model were used to examine the reliability and validity of CRCAT, and the data-model fit as well.

3.9.3 Data analysis

IBM SPSS (version 23) software was used for descriptive analysis for the quantitative data. Winsteps software 3.73.0 and Partial Credit Model (PCM) were used to estimating the reliability of the CRCAT by Cronbach alpha statistics, examining the construct validity of the clinical reasoning competency assessment tool (CRCAT), evaluating the nursing students' clinical reasoning competency, examining the items difficulty, estimating the fits between the items and the persons as well as finding out the problematic items of the test.

3.10 Conclusion

The study had two phases, phase one was the development of clinical reasoning competency assessment tool (CRCAT) using the four building blocks conceptual model suggested by Wilson (2005) including construct map, item design, outcome space and measurement. Phase two was the establishment of the validity and reliability of the developed clinical reasoning competency assessment tool (CRCAT) using test re-test reliability. ANOVA was used for groups comparison. Rasch measurement model analysis was used to examine the reliability and validity of CRCAT, and the data-model fit. The details of the development of CRCAT and the establishment of content validity and understandability was illustrated in the following Chapter Four.

CHAPTER FOUR

STUDY RESULTS OF PHASE ONE

4.1 Introduction

This chapter described the construction of the clinical reasoning competency assessment tool (CRCAT) using the conceptual framework of development model for measurement suggested by Wilson (2005), and the establishment of the content validity and understandability.

4.2 Construction of CRCAT

The clinical reasoning competency assessment tool (CRCAT) was constructed based on the conceptual framework of development model for measurement suggested by Wilson (2005), which consisted of four building blocks including construct map, item design, outcome space and measurement model (Figure 4-1). The construction of CRCAT using the four building blocks was described in detail in the following sections.

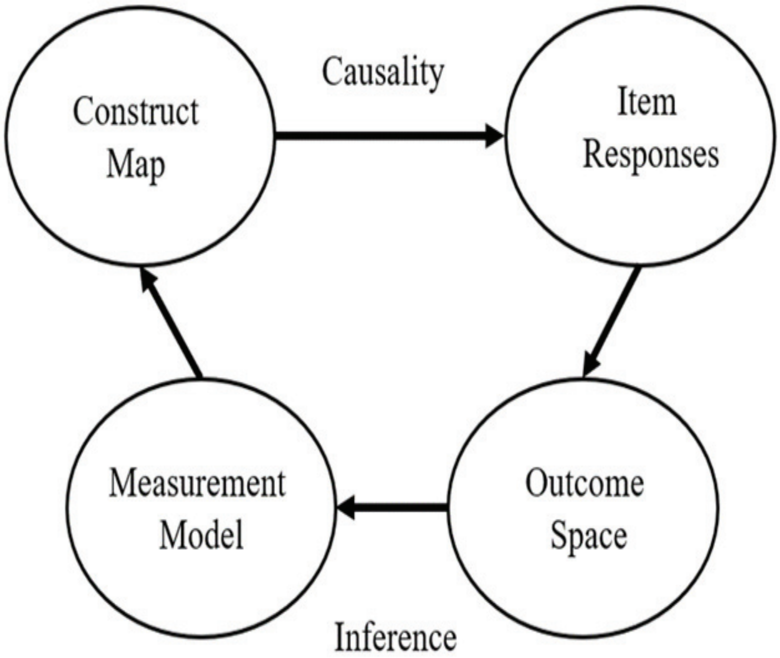

Figure 4-1: Conceptual framework of four building blocks (Wilson, 2005)

4.2.1 Building block one: construct map in construct "clinical reasoning competency"

As mentioned in chapter 2, clinical reasoning was a dynamic cognitive process having six levels of the cognitive process as classified by Benjamin Bloom and his team, they involved in clinical thought as well as clinical performance with increasing in complexity (Nkanginieme, 1997). The revised Bloom's cognitive process which was from remembering level to evaluating level was used as a continuum of clinical reasoning competency to develop the construct map (Figure 4-2) for the clinical reasoning competency assessment tool (CRCAT) in this study. The highest level of creating was not included in the continuum of clinical reasoining competency because of the considerations including: 1. The items of CRCAT were the clinical nursing procedures with focus on the cognitive process during the implementation of the nursing procedures, 2. The required level of clinical competence for the Enrolled Nursing students, 3. There should be differences among the individual created ideas or products making the scoring method difficult, 4. It was impossible and unrealistic that all clinical scenarios could stimulate innovative thoughts or create new products, 5. If the creating level was used in MCQ format, the products created certainly could not reflect the creating ability of the participants because the choice at creating level was decided by the test designer, and the choice only reflected the creating ability of the test designer not the test taker.

The variable being measured was “clinical reasoning competency”. The arrow running up and down the middle of the map indicated the continuum of the construct, running from “lowest” to “highest”. The left-hand side indicated qualitatively distinct groups of respondents; whereas, the right-hand side indicated qualitative differences in item responses, and the most right-hand side indicated the level of competency based on the level of cognitive complexity.

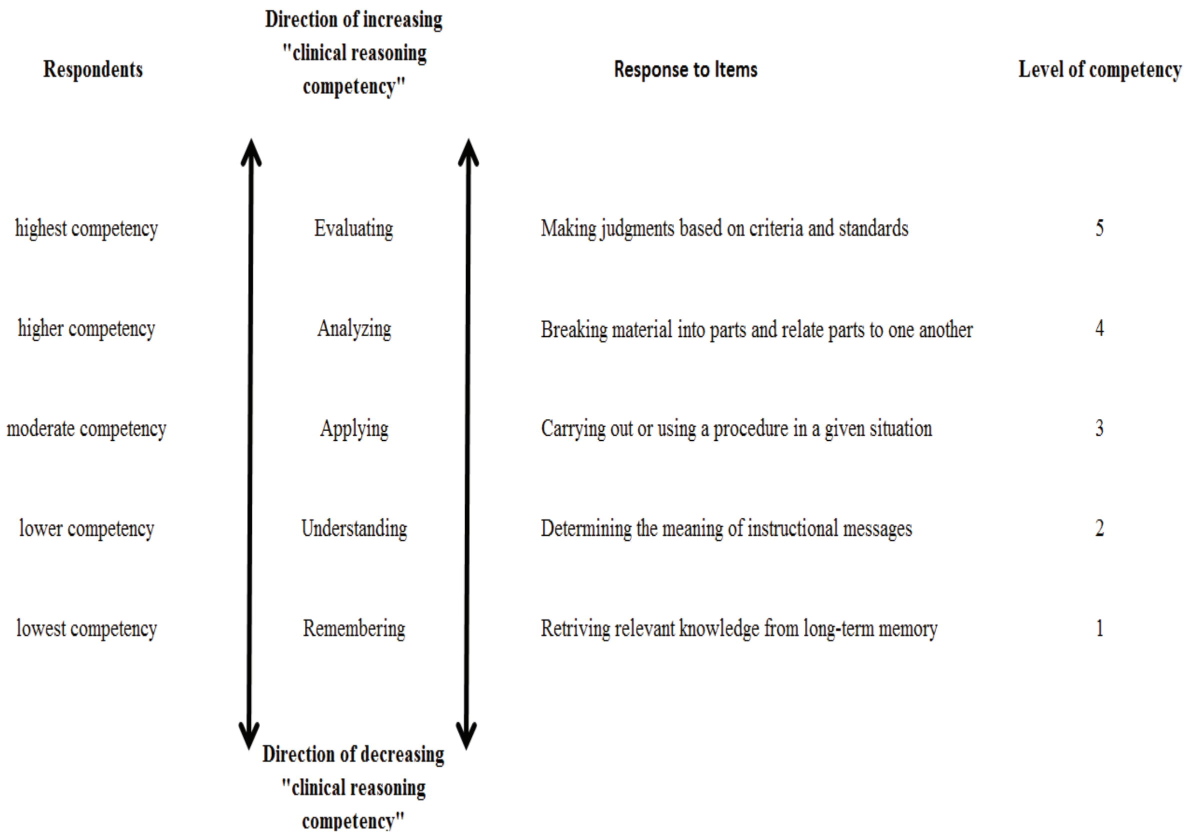

Figure 4-2: A Construct Map in Construct “Clinical Reasoning Competency”

4.2.2 Building block two: items design

Script was a cognitive sciences concept which explained how people had understood events that happened in the real world. Script activation was automatic and used in a strategic way to confirm hypotheses. The activated scripts had served to guide information selection, memorization and interpretation; hence, script-based clinical reasoning was very efficient (Charlin, Boshuizen, Custers and Feltovich, 2007). The script design used in the script concordance test (SCT) was based on the principle of the multiple judgments made in the clinical reasoning processes providing an assessment tool for clinical reasoning (Fournier, Demeester and Charlin, 2008). Therefore, in this study, according to the Association for Medical Education in Europe (AMEE) Guide No.75 (Lubarsky, Dory, Duggan, Gagnon, and Charlin 2013), the scripts of clinical situations with uncertainty was collected for construction of the items of CRCAT. The following was the process of items construction:

a. Constructing items

Fournier et al. (2008) suggested adopting three key features approach to constructing an item: 1. Recording a common clinical situation which was recently encountered in clinical practice. 2. Indicating relevant diagnostic hypotheses or management options that would be considered in the situation. This step provided the content for column 1 ('If you were considering:') and was designed to trigger the activation of specific illness scripts in the examinee's mind. 3. Indicating what clinical data might help to come to an appropriate decision or course of action in the situation, and what information would have little or no effect on examinee's reasoning. This step provided the content for column 2 ('And then you find:') and stimulated a data-gathering process. In this study, some scripts of clinical scenarios with situations of uncertainty that originating from everyday clinical experiences in medical and surgical wards were recorded and collected by the researcher. Twenty out of forty-nine scripts of clinical scenarios were selected, the scripts of clinical situations without uncertainty were not chosen. Each item of scenario was accompanied by three questions. Each question consisted of three columns as suggested by Fournier et al. (2008). Example 1 below was the first scenario in the CRCAT accompanied by 3 questions, each question associated with a particular case and was independent of the other two questions in the set. The first column indicated relevant management options in the item situation that provided the content for this column: "If you are considering of…". The second column indicated clinical data that helped making appropriate decision or taking a course of action in the item situation for this column: "And then you find…". The content of the third column : "Your consideration becomes…" was expected to elicit a range of positive ("+2" or "+1"), negative ("-2" or "-1") and neutral ("0") responses on the five-point Likert scale that for the persons' responses.

Example 1: (scenario one with three questions)

Scenario one: A man aged 45 presents with burning abdominal pain was admitted for investigation.

If you are considering of	And then you find	Your consideration becomes				
Q1: peptic ulcer	Patient has chest pain	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: insertion of NG tube	Patient has dark stools	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: prop up patient	Patient has vomiting	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Almost unnecessary
-1	Not useful
0	Nor less nor more useful
+1	Useful
+2	Absolutely necessary

b. Enhancing authenticity

All the twenty item scenarios of CRCAT were the representative cases observed and collected from both the medical wards and the surgical wards.

c. Number of item scenarios and questions

There were twenty item scenarios with a total of 60 questions in the clinical reasoning competency assessment tool (CRCAT).

4.2.3 The outcome space

In example 2, question 1 of scenario one, there were five choices to select, different choice represented different level of competency (Table 4-1).

Example 2: question 1 of scenario one

Scenario one: A man aged 45 presents with burning abdominal pain was admitted for investigation.

If you are considering of	And then you find	Your consideration becomes				
Q1: peptic ulcer	Patient has chest pain	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Almost unnecessary
-1	Not useful
0	Nor less nor more useful
+1	Useful
+2	Absolutely necessary

Table 4-1: Scale descriptions and competency levels of the 5-pts Likert scale (scenario one):

Q1	5-pts Likert scale	Scale Descriptions	Competency Level
	-2	Almost unnecessary	2
	-1	Not useful	3
	0	Nor less nor more useful	1
	+1	Useful	4
	+2	Absolutely necessary	5

In table 4-1 and Table 4-2, choice "0" (nor less nor more useful) indicates that the test taker was just recalling the related knowledge, not yet decided what to do. It implies that the stage of recalling in cognitive process was reached and the competency was at level 1. Choice "-2" (almost unnecessary) indicates that the test taker was Inferring which part of alimentary system would likely be affected, just understood the affected site was in GI tract but still not sure in which part of GI tract. It implies that the stage of understanding in cognitive process was reached and the competency was at level 2. Choice "-1" (not useful) indicates that the test taker was executing the brain activity thinking that what happened inside the stomach. It implies that the stage of applying in cognitive process was reached and the competency was at level 3. Choice "+1" (useful) indicates that the test taker was differentiating the differences and relationship between chest pain and burning abdominal pain. It implies that the stage of analyzing in cognitive process was reached and the competency was at level 4. Choice "+2" (absolutely necessary) indicates that the test taker was checking whether there was any signs relating to the heart problems apart from chest pain and critiquing how could the patient have the symptom of burning abdominal pain if it was the problem of heart or chest. It implies that the stage of evaluating in cognitive process was reached and the competency was at level 5.

Table 4-2: Illustration of five competency levels of clinical reasoning with scenario one

Category	Cognitive Process	Example Scenario one	Competency Level
Remembering	Recognizing	NA	1
	Recalling	Recall the pathophysiology of alimentary system	
understanding	Interpreting	NA	2
	exemplifying	NA	
	Classifying	NA	
	Summarizing	NA	
	Inferring	Infer which part of alimentary system would likely be affected	
Applying	Executing	What happened to the stomach?	3
	Implementing	NA	
Analyzing	Differentiating	Distinguish the difference between chest pain and burning abdominal pain	4
	Organizing	NA	
	Attributing	NA	
Evaluating	Checking	Are there any signs relating to the heart problems except chest pain?	5
	Critiquing	How could patient only have symptom of burning abdominal pain if it was the problem of heart or lung?	

The scoring method for the clinical reasoning competency was based on the Bloom's revised educational objectives taxonomy from the level of remembering to the level of evaluating. Score 1 was given to the chosen answer of lowest level of remembering, and score 2 was given to the chosen answer of second level of understanding, and score 3 given to the chosen answer of third level of applying, and score 4 given to the fourth level of analyzing, and the score 5 was given to the fifth level of evaluating.

4.2.4 The measurement model

a. Partial Credit Model

The Partial Credit Model (PCM), which has been developed by Masters, G. N. in 1982, was used in this study as a measurement model. It was presented as a straightly forward and logical application of Rasch's dichotomous model to a sequence of ordered response alternatives (Masters, 1988). Therefore, it could be used for the tests that containing polytomous items, such as Likert items and essays. The following equation was inferred the basic ideas of development of the partial credit model (WANG, 2010):

Let P_{nij} and $P_{ni(j\text{-}1)}$ denote the probability of scoring *j* and *j*-i on item *i* for person *n* respectively, θ_n denote the person *n*'s ability and δ_{ij} denote the *j*-th step difficulty of item *i*.

So, under the partial credit model, it is assumed:

$$\text{Logit}_{nij} \equiv \log[P_{nij} / P_{ni\,(j\text{-}1)}] = \theta_n - \delta_{ij} = \theta_n - (\delta_j + \tau_{ij})$$

δj is the mean of the step difficulties om item i and is called the overall difficulty. τij is the j-th deviation from the mean and is called the j-th threshold for the item i. Suppose item i has M + 1 categories and they are scored as 0, 1, …. M, then there will be M step difficulties of δij for that item. The first step difficulty describes how difficult it is by moving from category 1 (scoring 0) to category 2 (scoring 1), and so on, the M-th step difficulty describe how difficult it is by moving from category M to category M + 1. The step difficulties can be re-parameterized as a mean difficulty of δj and M thresholds ofτij. Thus, the partial credit model can be used for the tests with a sequence of ordered response alternatives (Masters,1988), such as questionnaires and survey with Likert scale, essay questions, etc.

b. Advantages of Partial Credit Model

The Rasch model was not a model developed only to fit data; rather it has been developed to diagnose data and to clean data in order to yield objective scales for persons and items (Wang, 2010). Moreover, the Item Response Theory (IRT) of Rasch model had several desirable advantages compared with the Classical Test Theory (CTT), they were including 1. The trait level estimate could be derived from any item for which properties were known, 2. Both item properties were directly linked to test behavior, and 3. The trait level and item properties were independent variables that could be estimated separately without additional data. Furthermore, since the construct of this study was clinical reasoning capability with different levels of competency, and varying degrees of credit has been assigned to students' attempts, such as by grading the students on a scale of five levels of competency based on their chosen answers. The partial credit model which was an extension of Rasch model could help to examine whether all the items in the developed clinical reasoning competence assessment tool (CRCAT) has been written appropriately as well as to diagnose noises in order to yield a meaningful and objective measurement assessment tool. Thus, partial credit model was the suitable measurement model being used to evaluate the model-data fit of the developed CRCAT.

4.3 Establishment of content validity and understandability of CRCAT

4.3.1 Content validity

For the content validity of the clinical reasoning competency assessment tool (CRCAT), method of expert review was used. The main aim was to determine whether the scenarios, the question items and the choices with different competency levels of each scenario were representative and relevant to the aim of the present study. Three nursing experts has been invited from clinical areas, they had over twenty years of clinical experience in both medical and surgical clinical areas. A standardized questionnaire was used to guide the review. Each of the experts was instructed to complete the questionnaire by assigning ratings on a 7-point Likert scale: strongly agree (7), agree (6), slightly agree (5), neutral (4), slightly disagree (3), disagree (2), strongly disagree (1). They were encouraged to provide written comments to justify their evaluations. Consensus was reached through a thorough discussion. The representative and relevancy of the scenarios, the question items as well as the scores that was assigned to different choices of competency levels of each scenario in the clinical reasoning competence assessment tool (CRCAT) were all confirmed.

4.3.2 Understandability

Five Enrolled Nursing students who were in other class were asked the three questions as tabulated in Table 4-3 after they has done the clinical reasoning competence assessment tool (CRCAT) in order to assess their understandability regarding the scenarios and the questions of the CRCAT. All of them replied that they understood the questions and knew what the scenarios had told and what the questions of each scenario had asked. They felt no difficulty in answering the questions and they were all feeling good after the test. Therefore, the understandability of clinical reasoning competency assessment tool (CRCAT) was confirmed.

Table 4-3: Understandability of the CRCAT questionnaire

Questions
1. Do you encounter any difficulties when filling in the questionnaire?
2. Do you understand the meaning of each question?
3. How do you feel after having completed the questionnaire?

4.4 Conclusion

The clinical reasoning competency assessment tool (Appendix A) for assessing the clinical reasoning competency of the Enrolled Nursing students was successfully constructed in phase one. The construction of item was based on the concepts of the comprehensive framework of four building blocks suggested by Wilson (2005) as well as the script format of association for medical education in Europe (AMEE) Guide No.75 (Lubarsky, Dory, Duggan, Gagnon and Charlin 2013). The first five levels of the cognitive process, which has been classified in the revised Bloom's Educational Objectives Taxonomy, was used as a grading ladder for the clinical reasoning competency. Moreover, the content validity as well as the understandability of CRCAT were also confirmed. Phase two was the evaluation of validity and reliability which were the two important technical features indicating whether an assessment tool was useful, suitable and with good quality so that it could be truly measuring what it was purported. The details of phase two was described in the following Chapter Five.

CHAPTER FIVE

STUDY RESULTS OF PHASE TWO

5.1 Introduction

This chapter described the process of implementation of clinical reasoning competency assessment tool (CRCAT) including data collection, data analysis using IBM SPSS (version 23) software for the descriptive analysis and groups comparisons. Winsteps 3.73.0 software for the Partial Credit Model (PCM) analysis. The establishment of reliability and validity of CRCAT were also elucidated.

5.2 Study Results

5.2.1. Demographic characteristics

The demographic distribution of the first year Enrolled Nursing students, final year Enrolled Nursing students who were in transition to Enrolled nurses, and the clinical nursing instructors were shown in Table 5-1 and Table 5-2:

Table 5-1: Demographics of nursing students

Demographic	Group 1 Final year students n = 41	Group 2 First year students n = 41
Range of age		
(18 – 25) years	35 (85.0%)	39 (95.0%)
(26 – 35) years	6 (15.0%)	2 (5.0%)
(36 – 45) years	0 (0.0%)	0 (0.0%)
(46 – 55) years	0 (0.0%)	0 (0.0%)
(56 – 65) years	0 (0.0%)	0 (0.0%)
Gender		
Female	34 (83.0%)	31 (76.0%)
Male	7 (17.0%)	10 (24.0%)
Education		
F. 5	0 (0.0%)	2 (4.8%)
F. 6 / F.7	27 (66.0%)	31 (75.6%)
Diploma	3 (7.3%)	2 (4.8%)
Higher Diploma	10 (24.0%)	6 (14.6%)
Bachelor	1 (2.4%)	0 (0.0%)
Master	0 (0.0%)	0 (0.0%)
Work experience		
No	24 (58.5%)	31 (75.6%)
Yes	17 (41.5%)	10 (24.4%)
Clinical experience		
CP 1 year	0 (0.0%)	41 (100%)
CP 2 years	41 (100%)	0 (0.0%)

There were 41 (100%) final year Enrolled Nursing students and 41 (100%) first year Enrolled Nursing students has participated in this study (Table 5-1). Among the Enrolled Nursing students, 83% and 76% were female, 17% and 24% were males respectively. Majority was in the age below 25. Over half of them including final year and first year Enrolled Nursing students (66% and 75.6%) has completed either Form 6 or Form 7 secondary school. Three (7.3%) and ten (24%) of the final year Enrolled Nursing students and two (4.8%) and six (14.6%) of first year Enrolled Nursing students have reached the education levels of Diploma and Higher Diploma respectively. One final year Enrolled Nursing student (2.4%) has owned bachelor's degree of study. 17 (41.5%) final year Enrolled Nursing students and 10 (24.4%) first year Enrolled Nursing students had working experience before entry into the Enrolled

Nursing training program; whereas, 24 (58.5%) final year Enrolled Nursing students and 31 (75.6%) year one Enrolled Nursing students were fresh graduates of secondary schools without any working experience. In table 5-2, 15 (100%) clinical nursing instructors has participated in this study. Among the participants, 73% was female, and 17% was male. Most of them aged from 36 to 55 year. Over half of them (53.3%) had master's degree of study. Six of them (40%) reached the Bachelor degree of study and one of them (6.7%) had degree of Higher Diploma. For the clinical experience, five of them (33.3%) had 6 years to 10 years of clinical experience as registered nurses, five of them (33.3%) had 11 years to 15 years and four of them were over 15 years of clinical experience as registered nurses. Six of them (40%) had less than 5 years of clinical experience as advanced practicing nurses, six of them (40%) had 6 years to 10 years of clinical experience as advanced practicing nurses. Three of them (20%) had over 20 years of clinical experience as nursing officers or advanced practicing nurses. Both the clinical nursing instructors and the Enrolled Nursing students had been working and practicing in the same hospitals in Hong Kong.

Table 5-2: Demographics of clinical nursing tutors

Demographic	Group 3 Clinical nursing instructors n = 15
Range of age	
(36 – 45) years	6 (40.0%)
(46 – 55) years	6 (40.0%)
(56 – 65) years	3 (20.0%)
Gender	
Female	11 (73.0%)
Male	4 (17.0%)
Education	
Higher Diploma	1 (06.7%)
Bachelor	6 (40.0%)
Master	8 (53.3%)
Clinical experience	
RN (0 – 5) years	1 (06.8%)
RN (6 – 10) years	5 (33.3%)
RN (11 – 15) years	5 (33.3%)
RN (16 – 20) years	2 (13.3%)
RN (21 – 25) years	2 (13.3%)
Clinical experience	
APN (0 – 5) years	6 (40.0%)
APN (6 – 10) years	6 (40.0%)
APN (21 – 25) years	3 (20.0%)

5.2.2 Correlations between demographics and clinical reasoning competency

The correlations between the demographics and the clinical reasoning competency of the participants including both the Enrolled Nursing students and the clinical nursing instructors were examined using Pearson product-moment correlation coefficient.

Table 5-3: Correlations between demographics and clinical reasoning competency

Demographic	Pearson's r correlation	Sig. (2-tailed)
Education level	-0.15	0.14
Clinical Practicum	-0.12	0.28
Clinical experience of clinical nursing instructors (years of RN)	0.06	0.83
Clinical experience of clinical nursing instructors (years of APN)	-0.21	0.44

In Table 5-3, there was no significant association found between the clinical reasoning competency and the demographic characters of the participants including education level (p = 0.14), clinical practicum (p = 0.28) and clinical experience of the clinical instructors (p = 0.83 and p = 0.44). Negative correlations were found between the clinical reasoning competency and the demographic characters including education level (r = -0.15), clinical practicum (r = -0.12) and clinical experience of the clinical nursing instructors (r = -0.21).

5.2.3 Test Re-test Reliability

Total 41 first year Enrolled Nursing students who were in other class finished the first attempt of CRCAT, and then repeated the CRCAT as 2nd attempt two weeks after the first attempt. Intra-class correlation coefficient was computed using person ability, and the results was shown in Table 5-4.

Table:5-4 Test Re-test Reliability of CRCAT

		95% Confidence Interval		F Test with True Value 0			
	Intraclass Correlation	Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	0.708	0.514	0.833	5.841	40	40	0.000
Average Measures	0.829	0.679	0.909	5.841	40	40	0.000

A high degree of reliability was found between test and retest measurements. The average measure Intraclass Correlation Coefficient (ICC) was 0.829 with a 95% confidence interval from 0.679 to 0.909〔F (40, 40) = 5.841, p<0.001〕.

5.2.4 Construct validity (correlation between PSI and CRCAT)

A Pearson product-moment correlation coefficient was computed with person ability for examining the relationship between problem solving ability using problem solving inventory (PSI) and the clinical reasoning competency using clinical reasoning competency assessment tool (CRCAT). Result showed a relatively weak but positive correlation between the problem solving ability and the clinical reasoning competency (r = 0.07, p = 0.52). It implied that increasing in problem solving ability was correlated with increasing in clinical reasoning competency.

5.2.5 Reliability and estimates

5.2.5.1 Person reliability

In Table 5-5, the person reliability using the Cronbach alpha coefficient (KR-20) was acceptable at 0.64, and the separation index was 1.34, the good separation index value was greater than the value of 2.0 (Bond & Fox, 2012, Nurulhuda et al., 2018). It implied that the clinical reasoning competency assessment tool (CRCAT) was less sensitive enough to distinguish the persons from the three different groups. The infit mean squares was at 1 and the outfit mean squares at 1 that were matched with the Rasch-modelled expectations of 1; whereas, the infit ZSTD was at zero and the outfit ZSTD at zero that were fit to the expected zero values implying that the CRCAT was high in both productivity and predictability for measurement.

Table 5-5: Summary of 97 measured persons

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	195.3	60	0.03	0.1	1	0	1	0
S.D.	17.1	0	0.17	0.01	0.16	1.1	0.18	1.1
MAX.	236	60	0.48	0.12	1.4	2.4	1.44	2.3
MIN.	155	60	-0.34	0.09	0.67	-2.6	0.64	-2.6

REAL RMSE 0.10 TRUE SD 0.13 SEPARATION 1.27 PERSON RELIABILITY 0.62
MODEL RMSE 0.10 TRUE SD 0.13 SEPARATION 1.34 PERSON RELIABILITY 0.64
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = 0.64

5.2.5.2 Item Reliability

In Table 5-6, the reliability of item difficulty estimates was high at 0.95, it indicated that the CRCAT had the ability to reproduce the hierarchy of items along the logits scale. It also suggested that this order of item hierarchy could be replicated with a high degree of probability if the items were given to other comparable cohorts (Bond & Fox, 20) .

Table 5-6: Summary of 60 measured items

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	315.8	97	0	0.08	1	0	1	0.1
S.D.	56.2	0	0.35	0.01	0.05	0.5	0.07	0.6
MAX.	436	97	0.77	0.1	1.18	2	1.2	2.2
MIN.	205	97	-1.08	0.06	0.88	-1.2	0.86	-1.4

REAL RMSE 0.08 TRUE SD 0.34 SEPARATION 4.31 ITEM RELIABILITY 0.95
MODEL RMSE 0.08 TRUE SD 0.34 SEPARATION 4.35 ITEM RELIABILITY 0.95

Also, the infit mean squares was at 1 and the outfit mean squares at 1 that were matched with the Rasch-modelled expectations of 1; whereas, the infit ZSTD was at zero and the outfit ZSTD at 0.1 very close to zero that were fit to the expected zero values implying that the CRCAT was high in both productivity and predictability for measurement. Moreover, the item separation recorded 4.35, the good separation index value was greater than the value of 2.0, it implied that the CRCAT had a good separation against the difficulty level according to the items (Bond & Fox, 2007; Nurulhuda et al., 2018).

5.2.6 Item performance

The item performance including item fit, item polarity, pathway map and Wright map were described in the following sections with Table 5-7, Table 5-8, Figure 5-1 and Figure 5-2 respectively.

5.2.6.1 Item fit

Item fit showed whether an item's response pattern had a good fit that reflecting the result of person and item measures was objective and interval (Wang, 2010). The item fit order of clinical reasoning competency assessment tool (CRCAT) was shown in Table 5-7.

Table 5-7: Item Fit Order

ENTRY NUMBER	MEASURE	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
4	**0.06**	**1.18**	**2**	**1.2**	**2.2**
11	-1.08	1.05	0.3	1.13	0.6
1	-0.15	1.08	0.9	1.12	1.3
57	-0.11	1.07	0.5	1.11	0.8
16	-0.35	1.05	0.4	1.11	0.6
13	-0.28	1.1	1.1	1.1	1.1
10	0.16	1.08	1.1	1.09	1.2
37	-0.34	1.06	0.5	1.09	0.6
18	-0.25	1.03	0.3	1.07	0.6
52	-0.01	1.05	0.4	1.06	0.5
42	0.19	1.05	0.4	1.06	0.4
51	-0.19	1.04	0.4	1.06	0.5
58	0.01	1.05	0.6	1.05	0.6
38	0.03	1.04	0.5	1.05	0.6
32	-0.15	1.05	0.4	1.04	0.3
23	-0.45	1	0.1	1.04	0.3
30	-0.4	1.04	0.3	1.02	0.2
9	0.25	1.04	0.3	1.02	0.2
24	0.24	1.03	0.3	1.03	0.4
17	0.65	1.03	0.3	1.03	0.3
12	0.07	1.02	0.3	1.03	0.3
48	0.3	1.02	0.2	1.03	0.3
59	-0.22	1	0	1.03	0.3
41	0.34	1.02	0.2	1.02	0.3
26	-0.07	1.01	0.2	1.02	0.2
2	0.5	1.02	0.2	1.02	0.2
20	-0.53	1	0.1	0.9	-0.4
33	0.34	0.99	-0.1	0.98	-0.2
55	0.18	0.99	-0.1	0.99	-0.1
15	-0.42	0.96	-0.2	0.99	0
50	0.39	0.97	-0.3	0.99	-0.1
29	-0.07	0.99	-0.1	0.98	-0.2
3	-0.65	0.98	-0.1	0.96	-0.1
34	0.01	0.95	-0.4	0.98	-0.1
5	-0.1	0.97	-0.2	0.98	-0.1
44	0.56	0.98	-0.2	0.96	-0.3
49	-0.49	0.98	-0.1	0.97	-0.1
8	-0.1	0.97	-0.2	0.97	-0.3
43	0.16	0.97	-0.3	0.97	-0.2
40	0	0.97	-0.3	0.96	-0.4
25	0.15	0.97	-0.4	0.95	-0.5
28	-0.09	0.95	-0.5	0.95	-0.5
31	-0.18	0.95	-0.3	0.92	-0.5
39	-0.29	0.95	-0.3	0.9	-0.6
27	0.2	0.94	-0.7	0.93	-0.8
35	-0.21	0.94	-0.5	0.94	-0.5
60	0.36	0.93	-0.8	0.93	-0.9
53	0.46	0.93	-1	0.93	-0.9
22	-0.44	0.92	-0.3	0.86	-0.6
46	-0.38	0.92	-0.4	0.88	-0.6
45	-0.04	0.92	-0.7	0.91	-0.8
54	0.77	0.88	-1.2	0.86	-1.4

The range of the infit mean squares was from 0.88 to 1.18 and that of the outfit mean squares was from 0.86 to 1.20 (Table 5-7). In Rasch Models, the value of the MNSQ outfit index was used to examine the conformity of the constructed items as well as to determine whether the items were developed appropriate to measure a latent construct. The MNSQ outfit index should be within the range between 0.6 to 1.4 if the items built were appropriate and suitable (Bond & Fox, 2007; Nurulhuda et al., 2018). From table 5-7, the MNSQ outfit index was within the range between 0.86 to 1.20 that was within the acceptable range between 0.6 to 1.4. It implied that the items of CRCAT were all constructed appropriately with good conformity.

5.2.6.2 Item polarity

According to Bond & Fox (2007) and Nurulhuda et al. (2018), the examination of item polarity was intended to test the extent of construction construct achieve, its aims and the relationship among the items that were built with respondents. In Table 5-8, the values shown on the PT-Measure Correlation were in the positive values except question item 4 and question item 13.

Table 5-8: Item polarity

ENTRY	INFIT		OUTFIT		PT-MEASURE	
NUMBER	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.
4	1.18	2	1.2	**2.2**	-0.12	0.24
13	1.1	1.1	1.1	1.1	-0.01	0.21
57	1.07	0.5	1.11	0.8	0.02	0.2
42	1.05	0.4	1.06	0.4	0.03	0.16
11	1.05	0.3	1.13	0.6	0.04	0.16
9	1.04	0.3	1.02	0.2	0.08	0.18
1	1.08	0.9	1.12	1.3	0.09	0.25
16	1.05	0.4	1.11	0.6	0.1	0.21
52	1.05	0.4	1.06	0.5	0.11	0.21
37	1.06	0.5	1.09	0.6	0.11	0.2
51	1.04	0.4	1.06	0.5	0.11	0.19
32	1.05	0.4	1.04	0.3	0.11	0.2
10	1.08	1.1	1.09	1.2	0.12	0.26
17	1.03	0.3	1.03	0.3	0.13	0.19
38	1.04	0.5	1.05	0.6	0.13	0.22
30	1.04	0.3	1.02	0.2	0.14	0.21
24	1.03	0.3	1.03	0.4	0.15	0.21
18	1.03	0.3	1.07	0.6	0.15	0.23
58	1.05	0.6	1.05	0.6	0.15	0.25
48	1.02	0.2	1.03	0.3	0.16	0.2
41	1.02	0.2	1.02	0.3	0.17	0.21
2	1.02	0.2	1.02	0.2	0.18	0.21
6	1	0	1.01	0.1	0.18	0.19
56	1	0.1	1	0.1	0.19	0.2
47	1	0	1	0	0.19	0.19
21	1	0	1	0	0.19	0.2
26	1.01	0.2	1.02	0.2	0.2	0.23
36	1.01	0.2	1.01	0.1	0.2	0.22
19	1.01	0.1	1	0	0.2	0.21
12	1.02	0.3	1.03	0.3	0.21	0.24
7	1.01	0.1	1.02	0.1	0.21	0.19
20	1	0.1	0.9	-0.4	0.21	0.18
23	1	0.1	1.04	0.3	0.21	0.2
59	1	0	1.03	0.3	0.22	0.22
55	0.99	-0.1	0.99	-0.1	0.23	0.2
14	1	0	1	0.1	0.24	0.24
33	0.99	-0.1	0.98	-0.2	0.24	0.21
3	0.98	-0.1	0.96	-0.1	0.24	0.18
44	0.98	-0.2	0.96	-0.3	0.25	0.2
29	0.99	-0.1	0.98	-0.2	0.25	0.21
49	0.98	-0.1	0.97	-0.1	0.26	0.17
5	0.97	-0.2	0.98	-0.1	0.27	0.2
15	0.96	-0.2	0.99	0	0.27	0.19
8	0.97	-0.2	0.97	-0.3	0.27	0.23
50	0.97	-0.3	0.99	-0.1	0.28	0.22
43	0.97	-0.3	0.97	-0.2	0.29	0.22
40	0.97	-0.3	0.96	-0.4	0.29	0.22
34	0.95	-0.4	0.98	-0.1	0.32	0.23
28	0.95	-0.5	0.95	-0.5	0.32	0.23
25	0.97	-0.4	0.95	-0.5	0.33	0.25
39	0.95	-0.3	0.9	-0.6	0.33	0.21
31	0.95	-0.3	0.92	-0.5	0.34	0.22
35	0.94	-0.5	0.94	-0.5	0.35	0.22
27	0.94	-0.7	0.93	-0.8	0.36	0.23
60	0.93	-0.8	0.93	-0.9	0.37	0.23
46	0.92	-0.4	0.88	-0.6	0.38	0.18
53	0.93	-1	0.93	-0.9	0.38	0.23
22	0.92	-0.3	0.86	-0.6	0.38	0.19
45	0.92	-0.7	0.91	-0.8	0.39	0.21
54	0.88	-1.2	0.86	-1.4	0.47	0.21

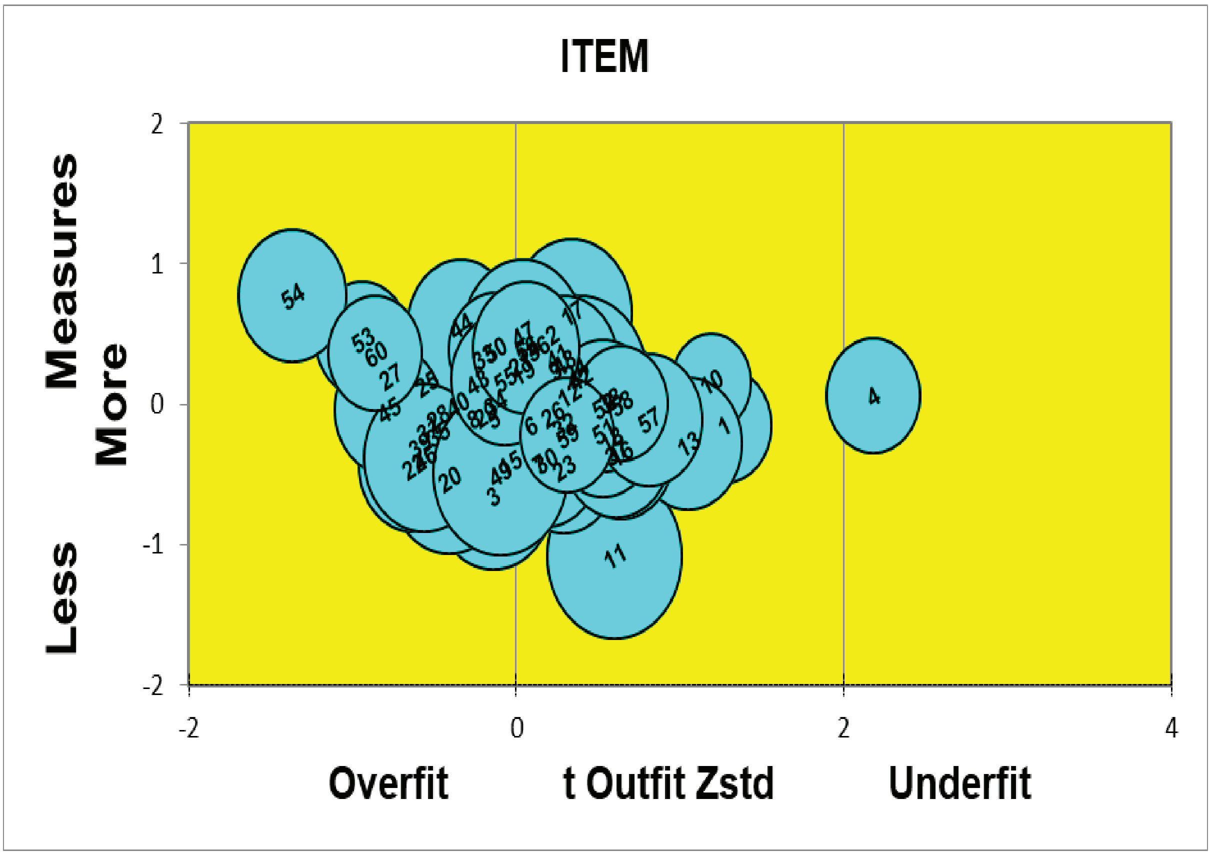

Figure 5-1: Item Pathway Map

5.2.6.3. Pathway map

In Figure 5-1, almost all the items fell between -2.0 and +2.0 except the question item 4, it seemed a bit erratic that was off the pathway to the right. Nevertheless, all items were located within a band of over the logits spreading around the zero origin from +1.5 to -2.0. The erratic behavior of item 4 was further investigated in the following chapter.

5.2.6.4 Wright map

The Wright map (Figure 5-2) was organized as two vertical histograms. The left side showed participants and the right side showed question items. The left side of map showed the distribution of the measured clinical reasoning competency of the participants from most able at the top to least able at the bottom. The question items on the right side of the map were distributed from the most difficult at the top to the least difficult at the bottom. On the left side, the Wright map showed the mean (M) and two standard deviation points (S=one SD and T=two SD) for measured participant clinical reasoning competency. On the right side of the map, the mean difficulty of the question items (M) and two standard deviation points (S=one SD and T=two SD) for the question items were shown (Bond & Fox, 2007). The Wright map showed that both the mean (M) clinical reasoning competency of the participants and the mean (M) difficulty of the question items were at the level of "0" logits.

```
MEASURE     PERSON - MAP - ITEM
                <more>|<rare>
    1                 +
                      |
                      |  I0054
                      |T I0017
                      |  I0044
                   .  |  I0002  I0047  I0053
                  .# T|S I0014  I0036  I0050  I0056  I0060
                  ##  |  I0009  I0021  I0033  I0041  I0048
              ###### S|  I0010  I0019  I0024  I0025  I0027  I0042  I0043  I0055
        .############ |  I0004  I0012
    0     ########## M+M I0034  I0038  I0040  I0045  I0052  I0058
            .######## S| I0001  I0005  I0008  I0026  I0028  I0029  I0032  I0057
               #####  |  I0006  I0018  I0031  I0035  I0051  I0059
                 .## T|  I0013  I0016  I0037  I0039
                      |S I0007  I0015  I0022  I0030  I0046
                      |  I0020  I0023  I0049
                      |  I0003
                      |T
                      |
                      |
   -1                 +
                      |  I0011
                      |
                      |
                      |
                      |
                      |
                      |
                      |
                      |
   -2                 +
                <less>|<frequent>
EACH "#" IS 2. EACH "." IS 1.
```

Figure 5-2: Item-person map

Each symbol of '#' and '.' on the left of the Figure 5-2 were representing two persons and one person respectively. The number on the right was the question items, sixty locations are calibrated with a default mean difficulty set at 0.0. The left-hand side of the Wright map was the locations of person performance that telling something about the clinical reasoning competency of the persons who had completed the clinical reasoning competence assessment tool (CRCAT). They were spreading out over about 2 logits that from logits -0.5 to the logits +0.5 with more persons gathering at the area from logits -0.25 to logits +0.25. By looking at the person distribution against the item distribution, most of the persons were located opposite the items at the middle of the scale around the logits from -0.5 to +0.5 where most of the items were located. Regarding the item difficulty, the items of 17, 44 and 54 were the toughest questions; whereas the items of 3, 7, 15, 20, 22, 23, 30, 46 and 49 are the easiest. The Wright map showed that many of the items were within plus or minus one standard deviation, there were also all participants aligned within one standard deviation. Therefore, this sample included sufficient numbers of items in the center of the item distribution that was close to differentiate persons at different levels of logits suggesting that the person distribution was matched to the items that measured their perceived competency.

5.2.7 Group comparison

A one-way between groups and within groups using analysis of variance (ANOVA) had been conducted to compare the clinical reasoning competency among the three groups. The analysis of variance (ANOVA) was a powerful analytic tool for analyzing and determining whether there were any statically significant differences between the means of three or more independent groups (Portney & Watkins, 2009).

Table 5-9: Comparison among three groups using ANOVA

Clinical competency	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	0.26	2	0.13	5.100	**0.01**
Within Groups	2.42	94	0.03		
Total	2.68	96			

In Table 5-9, there was a significant difference in clinical reasoning competency among the three groups at the p<0.05 level〔F (2, 94) = 5.100, p = 0.01〕.

The results of Post hoc comparisons using Tukey HSD test were shown in the following two tables (Table 5-10 & Table 5-11).

Table 5-10: Descriptive of clinical reasoning competency of three groups

Clinical Competency	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
clinical nursing instructors	15	-0.00	0.15	0.04	-0.09	0.08	-0.23	0.24
final year students	41	0.09	0.15	0.02	0.04	0.14	-0.14	0.48
first year students	41	-0.02	0.17	0.03	-0.07	0.04	-0.34	0.28
Total	97	0.03	0.17	0.02	-0.00	0.06	-0.34	0.48

Table 5-11: Multiple Comparisons of clinical reasoning competency among three groups

Dependent Variable: Clinical competency
Tukey HSD

(I) groups	(J) groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Clinical Nursing instructors	final year students	-0.09	0.05	0.13	-0.21	0.02
	first year students	0.01	0.05	0.96	-0.10	0.13
final year students	clinical nursing instructors	0.09	0.05	0.13	-0.02	0.21
	first year students	0.11	0.04	**0.01**	0.02	0.19
first year students	clinical nursing instructors	-0.01	0.05	0.96	-0.13	0.10
	final year students	-0.11	0.04	**0.01**	-0.19	-0.02

Post hoc comparisons using Tukey HSD test (Table 5-10 and Table 5-11) showed that there was significant difference in clinical reasoning competency ($p = 0.01$) between the first year Enrolled Nursing students and the final year Enrolled Nursing students. It implies that the clinical reasoning competency of the Enrolled Nursing students increases as their clinical experience increases. However, no significant difference was found between the clinical nursing instructors and the final year Enrolled Nursing students ($p = 0.13$), and between the clinical nursing instructors and the first year Enrolled Nursing students ($p = 0.96$).

5.3 Conclusion

A useful assessment tool should be truly measuring what it was purported to measure and had supportive outcomes (Portney & Watkins, 2009). Therefore, evaluating the validity and reliability of the newly developed assessment tool was very important. In this chapter, the reliability and validity of the clinical reasoning competency assessment tool (CRCAT) were tested and proved. The results of the data analysis using Rasch measurement model also reflected a fit model-data matched with person reliability of Cronbach alpha coefficient at acceptable level. The infit and outfit mean squares as well as the infit and outfit ZSTD were very close to the Rasch-modelled expectations of 1 and 0 respectively indicating that the CRCAT was high in both productivity and predictability for measurement. Moreover, all except question item 4 & question item 13 presented positive values of PT-Measure Correlation. The Wright map showed the person distribution was matched to the items that measured their perceived competency. Post hoc comparisons using Tukey HSD test showed a significant difference in clinical reasoning competency between the first year Enrolled Nursing students and the final year Enrolled Nursing students was found at the $p<0.05$ level. Therefore, it was concluded that the clinical reasoning competency assessment tool (CRCAT) was a reliable and valid clinical assessment tool used to evaluate the clinical reasoning competency for the Enrolled Nursing students. Discussion and implication on the study results of CRCAT were further explored and described in the following Chapter Six.

CHAPTER SIX

DISCUSSION AND IMPLICATION

6.1 Introduction

The usefulness of a measurement tool in research depends on the extent to which the data can be relied on as accurate and meaningful indicators of an attribute such as the clinical reasoning competency. The first prerequisite of a recognized measurement assessment tool was reliability that was the extent to which a measurement was consistent and free from error. The second prerequisite was validity that a measurement tool was measuring what it was intended to measure. Both reliability and validity were the two essential technical features to be used to determine whether a new assessment tool was useful (Portney & Watkins, 2009). Therefore, different general approaches to testing of reliability and validity of the clinical reasoning assessment tool (CRCAT) were used in this study, moreover, the Rasch partial credit model analysis techniques were used to evaluate the fitness of the model and data such as person ability, item difficulty and the fits between the persons and the items, and the results were described in previous Chapter Five. In this chapter, the study results are discussed in the following sections in responses to the research questions and the hypotheses mentioned in Chapter One.

6.2 Research question 1: What are the reliability and validity of the clinical reasoning competency assessment tool (CRCAT)?

6.2.1 Reliability of clinical reasoning competency assessment tool (CRCAT)

Reliability refers to the consistency of a measurement tool. There are three types of consistency: first type is over time named test-retest reliability, second type is across items named internal consistency and the third type is across different researchers named inter-rater reliability (Jackson, 2012). In this study, test-retest reliability and internal consistency were evaluated for the consistency over time and the consistency across items accordingly; however, the inter-rater reliability tests was not included since there was only one researcher for this study. Test-Retest reliability is usually analyzed using intra-class correlation coefficient (ICC) as it can reflect both correlation and agreement (Portney & Watkins, 2009), also, ICC is a widely used reliability index, it helps readers to better understand their own clinical practice (Koo & Li, 2016). Since ICC reflects not only the degree of correlation but also agreement between measurements, ICC is a more desirable measure of reliability compared with those non-ideal measures, such as paired t test and Bland-Altman plot were for the agreement only (Koo & Li, 2016). According to Trevethan (2016), there are three models of intra-class correlation coefficients (ICCs), each of the models differed is in terms of where the sources of statistical variability are believed to be, and these sources of variability includes the participants in study. Model 1 is used when a range of different raters assess different participants, and there is no match between raters and participants. This situation is infrequent. This model will usually produce lower ICCs than do the other two models. Model 2 is used when the same raters assess all participants, and theoretically the raters are regarded as being randomly selected, as are the particular participants. This model is particularly appropriate when different raters' consistency in using a particular instrument is being assessed. Compared with model 1 and model 2, model 3 is the most appropriate model to be used to assess whether the specific raters in their specific study are consistent, either within themselves (intrarater reliability) or between each other (interrater reliability), and the participants are as random, further, model 3 can apply to many situations, including those used to assess intrarater reliability and test-retest reliability. It will usually produce the highest ICCs (Trevethan, 2016; Brozek & Alexander, 1947; Muller & Buttner 1994). Therefore, in this study, the test-retest reliability was analyzed using model 3, average measures and consistency to testing intra-class correlation coefficient (ICC).

There are two most common sets of criteria for interpreting results of ICC. Fleiss categories regard all ICCs above 0.75 as excellent, whereas Portney and Watkins characterize ICCs from 0.75 up to 0.90 more reservedly as merely good, and they avoid use of the word excellent altogether (Trevethan, 2016). In table 5-4, result shows that the Intraclass Correlation Coefficient (ICC) was 0.829 with 95% confident interval of an ICC estimate was between 0.679 and 0.909. According to the above two most common sets of criteria for interpretation of ICCs, the results indicate that the clinical reasoning assessment tool (CRCAT) had good to excellent test-retest reliability.

6.2.2 Validity of Clinical Reasoning Competency Assessment (CRCA) Tool

According to Portney & Watkins (2009), measurement of validity is the extent to which the new instrument measures what it is intended to measure, it is usually devised for purposes of discrimination, evaluation, or prediction. Validity implies that a measurement is relatively free from error, it is evaluated within the context of the new instrument's intended use. A valid test is a reliable instrument used to make inferences about the magnitude of a particularly variable based on a relevant observable behavior or response. Validation is a process of hypothesis testing, determining if the test scores are related to specific behaviors, characteristics or levels of performance. Evidence to support hypothesis is generally defined according to four types of validity including face validity, content validity, criterion-related validity and construct validity. Since the face validity is the weakest form of measurement validity, it was not used as a test for the validity of CRCAT. The establishment of content validity and understandability for the CRCAT was already described in Chapter Three. As aforesaid, both the problem solving inventory (PSI) and the clinical reasoning competency assessment tool (CRCAT) are the tools focusing on assessing problem solving related ability but the former one is perception-focused; whereas, the focus of the latter one is cognitive process. Since there was no similar clinical assessment tool constructed using script-based scenarios and revised Bloom's cognitive process as capability of competency, both the problem solving inventory (PSI) and the clinical reasoning competency assessment tool (CRCAT) were used to test the construct validity of the developed CRCAT in this study.

For the construct validity of the CRCAT, Pearson product-moment correlation coefficient was computed with person ability to testing the relationship between the problem solving ability using problem solving inventory (PSI) and the clinical reasoning competency using the clinical reasoning competency assessment tool (CRCAT). Results showed that the Pearson's r was not significant ($r = 0.07$, $p = 0.52$). Analytically, there are several ways used to interpret the values of Pearson's r, and the most commonly agreed interpretation is that a Pearson correlation coefficient of < 0.1 indicates a negligible or weak and >0.9 a very strong relationship; however, correlations were frequently misunderstood such as a correlation coefficient which was close to zero demonstrating the variables were not related. In fact, correlations did not describe the strength of agreement between two variables that could exhibit a high degree of correlation but could at the same time disagree substantially. Also, different relationship between variables could result in similar correlation coefficients (Schober, 2018). Therefore, although the Pearson's r was 0.07 that is interpreted as a negligible or weak value, it may not really reflect any relationship between PSI and CRCAT; however, the positive Pearson's implies the two variables were going towards the same direction. Moreover, there were three constructs including problem-solving confidence, avoidance style and personal control were measured by the PSI, which is multidimensionality ; whereas, the CRCAT was unidimensionality measuring one construct of clinical reasoning competency only. Thus, this may be the cause leading to a relatively low value of Pearson's r between PSI and CRCAT.

6.3 Research question 2: What is the Partial Credit Model-data fit of the clinical reasoning competency assessment tool (CRCAT)?

Partial credit model is the extension of Rasch model developed by Wright & Master in 1982. This model specifically incorporates the possibility of having differing numbers of response opportunities for different choices on the same question item, each choice has its own rating credit, the multiple-choice question test where the responses that are incorrect but indicate some knowledge or meaning will be given partial credit towards a correct response, the amount of partial correctness varies across the choices of the question item (Bond & Fox, 2007). As mentioned in Chapter Four, the scoring method for the clinical reasoning competency was based on Bloom's revised educational objectives taxonomy from the level of remembering to the level of evaluating. Score 1 given to the chosen answer of lowest level of remembering, and score 2 given to the chosen answer of second level of understanding, and score 3 given to the chosen answer of third level of applying, score 4 given to the fourth level of analyzing and score 5 to the fifth level of evaluating (Table 4-1 & Table 4-2). The Rasch partial credit model analysis has several advantages when measuring students' competencies (Eggert & Bogeholz, 2009), such as 1. Transformation of raw scores to logarithmic units (logits) making the analysis of students' abilities and items difficulties possible on an equal interval linear scale. 2. Presentation of both persons and items with their abilities and difficulties respectively was plotted on the same logit scale (Wright maps), more able persons and more difficult items were at the top of the map; whereas, the less able person and less difficult items were at the base of the map that providing a very useful base to evaluate the functioning of a test or an assessment tool. 3. Reliability indices used to calculate both for persons and items, namely person separation reliability and item separation reliability, the former one was an estimate of how well one could differentiate persons on the measured variable and the latter one was an estimate of how well the sample of subjects had spread out the items along the measure of the instrument (Fisher, 1992; Eggert & Bogeholz, 2009). 4. Fit statistics could highlight unexpected response behaviors that contribute to the test or instruments' validity. Therefore, the Rasch partial credit model analysis was used in this study.

In Rasch measurement, the concept of 'fit' is as a quality-control mechanism, and the fit statistics are to reflect whether the assumption of unidimensionality can be held up empirically and to determine whether the item estimations being held as meaningful quantitative summaries of observations. Fit as a quality-control mechanism, the fit values are on a standardized *t* scale horizontally with acceptable values falling between -2.0 and +2.0 for the study sample sizes between 30 and 300 (Bond & Fox, 2007). The sample size of the present study is 97 that is within the said range. According to the item pathway shown in Figure 5-1, almost all the items fall between -2.0 and +2.0 except the item 4, it seems a bit erratic that is off the pathway to the right; nevertheless, all items are located within a band of over the logits spreading around the zero origin from +1.5 to -2.0. The erratic behavior of item 4 will be further investigated in the following section.

6.3.2 Items Fit

Items fit measures the constructs that can be seen through the infit and outfit Mean Square (MNSQ), and the outfit and infit MNSQ shall be in the range of 0.6 to 1.4 to ensure the items are suitable for measuring the constructs (Yunus et al., 2015; Bond and Fox, 2007). The fit statistics are used to determine how well any set of empirical data meets the requirements of Rasch model, the Infit and outfit statistics are reported as mean squares (MNSQ) with an expected value of +1 and a range from 0 to positive infinity. If an infit or outfit mean square value is greater than 1, such as 1.2, it indicates 20% more variation in the observed data than the Rasch model predicts; if an outfit mean square value is less than 1, such as 0.78, it indicates 22% less variation in the observed response pattern than is modeled. Infit and outfit mean square values are always positive and are used to monitor the compatibility of the data with the model used. Infit and outfit statistics are also reported as a standardized form (ZSTD) with expected value of zero. It is interpreted as the observed data having less compatibility with the model than expected when the infit and outfit z standardized values greater than +2 or less than -2, the positive or negative values indicate more variation or less variation respectively than modeled (Bond and Fox, 2007).

The general interpretation of fit statistics is if mean square (MNSQ) is greater than 1.3 and z standardized value (ZSTD) greater than 2.0, it is an underfitting performance and is interpreted as unpredictable item performance with too haphazard response pattern and too much variation which should be reflected on what is going wrong; if mean square (MNSQ) is lower than 0.75 and z standardized value (ZSTD) lower than -2.0, it is then an overfitting performance and is interpreted as Guttman with too determined response pattern and too little variation which may lead to wrongly thinking that the quality of the measures is better than it actually is (Bond and Fox, 2007). In table 5-2, the mean squares (MNSQ) infit and outfit statistics of all items are near to the expected value of 1 and within the acceptable range between 0.75 and 1.3, the infit MNSQ is from 0.88 to 1.18, and the outfit MNSQ is from 0.86 to 1.2. The z standardized (ZSTD) statistics of all items except item 4 are within the acceptable range between -2.0 and +2.0, the infit z standardized (ZSTD) are from -1.2 to 2.0, and the outfit z standardized (ZSTD) are from -1.4 to 2.2. The outfit z standardized (ZSTD) of erratic item 4 is 2.2 which is greater than the expected value; however, if the outfit and infit MNSQ are in the range of 0.6 to 1.4, the ZSTD index can be ignored (Yunus et al., 2015; Linacre, 2007). From Table 5-2, both the outfit and infit MNSQ of item 4 are within the range of 0.6 to 1.4, as aforementioned, in Rasch models, the outfit index is used to examine the conformity of the constructed item and to determine whether the item is built appropriately and suitably, the outfit and infit MNSQ which are in the range of 0.6 to 1.4 are used to ensure the items are suitable for measuring the constructs (Yasin et al. 2015; Bond and Fox 2007). Since the outfit MNSQ and infit MNSQ of item 4 are 1.2 and 1.18 respectively, item 4 can remain in the CRCAT. In conclusion, both the Figure 5-1 Pathway map and the Table 5-7 Items Fit Order show model-data fit.

6.3.3 The Wright Map

The Wright map, also called person-item map, provided a very useful and easy observable equal linear logit scale to evaluating the functioning of an assessment tool or an instrument (Eggert & Bogeholz, 2009).

In Figure 5-2, on the left side, the Wright map shows the mean (M) and two standard deviation points (S=one SD and T=two SD) for measured person clinical reasoning competency; on the right side of the map, the mean difficulty of the question items (M) and two standard deviation points (S=one SD and T=two SD) for the question items are shown (Bond & Fox, 2012). The Wright map shows that both the mean (M) clinical reasoning competency of persons and the mean (M) difficulty of the question items are at the level of "0" logits. Each symbol of " # " and " . " on the left represents two persons and one person respectively. Persons including nursing students and clinical nursing instructors are plotted on the left side of the linear logit scale as a function of their clinical reasoning competency; whereas question items are plotted on the right side of the same logit scale as a function of items difficulty. More competent persons and more difficult question items are at the top of the map, less competent persons and less difficult question items are at the base of the map. Persons who are plotted on the same level as the question items have a 50% chance of getting that question items correct. In Figure 5-2, Persons are spreading out between logits -0.5 to logits +0.5, most of them are located opposite the question items at around the logits from -0.25 to +0.25. The mean person estimate is located at the zero of logits. Although there is sizable group of question items (2, 17, 44, 47, 53, 54) located at above the logits 0.5 left without targeted persons, and so are the easy question items (3, 7, 11, 15, 20, 22, 23, 30, 46, 49) located under the logits -0.5, those question items can be remained in the assessment tool for the future study use for other different sample subjects. Figure 5-2 person-item map shows model-data fit indicating that the clinical reasoning competency assessment tool (CRCAT) is a well-matched assessment tool to the sample of subjects.

6.3.4 Person Reliability and Item Reliability

Internal Consistency (Homogeneity) can be assessed using Cronbach Alpha Coefficient, which reflects the extent to which all the items in a test measure the same concepts or construct such as if the test is used to test students' knowledge of research design, the items should reflect a summary of that knowledge; and the test should not include items on health policy (Portney & Watkins, 2009). Alpha was developed by Lee Cronbach in 1951. It is the most widely used objective measure of reliability, it provides a measure of the internal consistency of a test or an instrument, and is connected to the inter-relatedness of the items within the test or instrument, it can also be used to confirm whether the sample of items is actually unidimensional (Tavakol & Dennick, 2011). In Rasch measurement model, it provides indices that can help the researcher to determine whether there are enough items spreading along the continuum with difficulty levels, also enough spread of ability among persons (Bond & Fox, 2007). The person reliability index indicates the replicability of person ability ordering if the same persons are given another parallel set of items measuring the same construct; whereas, the item reliability index indicates the replicability of items placements along the pathway if the same items are given to another group of persons with same size behaving the same way (Wright & Masters, 1982; Bond & Fox, 2012). Table 5-5 shows that both the Cronbach Alpha Coefficient and the person reliability are acceptable at 0.64; however, the person separation is 1.34 which is lower than the good separation index indicating that the replicability of person ordering cannot be as expected if the same sample of persons are given another parallel set of items measuring the same construct (Bond and Fox, 2007); in other words, the number of targeted items are not enough spread of ability among the persons, and more targeted items should be developed.

For Cronbach's alpha test of reliability, there are a wide range of different qualitative descriptors used by different authors to interpret its values, such as excellent (0.93-0.94), strong(0.91-0.91), reliable(0.84-0.90), robust(0.81), fairly high(0.76-0.95), high(0.73-0.95), good(0.71-0.91, relatively high(0.70-0.77), slightly low(0.68), reasonable(0.67-0.87), adequate(0.64-0.85), moderate(0.61-0.65), satisfactory(0.58-0.97), acceptable(0.45-0.98), sufficient(0.45-0.96), not satisfactory(0.4-0.55) and low(0.11). Since, as shown in Table 5-5, the Cronbach's alpha of the CRACT was at 0.64, it could be interpreted as adequate, moderate, sufficient, acceptable and satisfactory according to the above descriptors being used by different researchers. Cronbach had suggested that a high value of alpha was desirable if the assessment tool was used to assign a score to a person; but he argued that the key point should be that the scores obtained when using an instrument or assessment tool had to be interpretable, this was often possible without needing very high values of alpha, though increasing items could probably increase the alpha value (Taber, 2016). Some assessment tools with low values of alpha could still be proved useful (Schmitt, 1996; Taber, 2016).

In Table 5-6, the item reliability index is 0.95 implying that the CRCAT has the ability of reproducing the hierarchy of items along the logits scale, and the replicability of items is high, and if the items are given to another sample of persons with same size, the outcome behaviors detected will be the same way. Although the items, which located at the top difficult level and the bottom easy level, are without targeted persons can be remained in the CRCAT for future study using different sample of subjects.

Since the number of question items has profound effects on alpha value (Cortina, 1993) and the high alpha value is desirable for the CRCAT as it was used to assign different scores to students at different levels of clinical reasoning competency in this study. Hence, it was necessary to increase the number of scenarios from each specialty including medical and surgical clinical areas increasing the number up to, say, one hundred and twenty question items. By doing so, it would certainly help to improve the internal reliability of the clinical reasoning competency assessment tool (CRCAT).

6.3.5 Item Polarity

In table 5-8, analysis on point measure correlation (PMC) indicates the construct validity of the question items. All question items were positive (> 0) except question item 4 & question item 13. The infit and outfit mean square MNSQ of question item 4 are 1.18 and 1.2 respectively, question item 13 are 1.1 and 1.1 respectively. They are within range (0.6 to 1.4). The infit and outfit Z-standard value ZSTD of question item 4 are 2.0 and 2.2, question item 13 are 1.1 and 1.1. Only the outfit ZSTD of question item 4 is out of range (-2 to +2). Since only one control could not be met by question item 13 and two controls by question item 4, they could not be considered as misfit. Item is considered as misfit only when all three controls were out of range, also, guessing is a measurement disturbance, sometimes makes the good item to bad. Misfit does not mean "throw it away"; it means "find out why". (Bond & fox, 2015), therefore, question item 4 and question item 13 can remain in CRCAT but needed to be investigated and revised as these two question items may be difficult to answer by the respondents.

6.4 Research question 3: Are there significant differences in clinical reasoning competency between the final year Enrolled Nursing students and the clinical nursing instructors, and between the final year Enrolled Nursing students and the first year Enrolled Nursing students?

6.4.1 Clinical reasoning competency among groups

The original ideas of analysis of variance (ANOVA) has been developed by the English Statistician Sir Ronald A. Fisher (1890-1962). ANOVA is a statistical procedure concerned with comparing means of several samples. Nowadays, it is the commonly used advanced research methods among different professionals. The powerful statistical techniques of one-way ANOVA are usually used to analyze variability in data in order to infer the inequality among population means (Ostertagova & Ostertag, 2013). As the ANOVA is based on the same assumption with the t-test and is the appropriate method for a comparison of more than two groups (Kim, 2014). Therefore, in present study, ANOVA was conducted for the comparison between groups and within groups to evaluate the clinical reasoning competency among the three groups including the final year Enrolled Nursing students, the first year Enrolled Nursing students and the clinical nursing instructors. In table 6-1, there was a significant difference in clinical reasoning competency among the three groups at the $p<0.05$ level〔$F(2, 94) = 5.100$, $p = 0.01$〕.

6.4.2 Multiple Comparison of Clinical reasoning competency among three groups

As the significant difference has been detected by the overall F test as shown in Table 5-9, multiple comparison of clinical reasoning competency using Tukey's HSD test was used to examine what specific pair of group means showed difference and what pair of group means did not. In Table 5-10, the means of clinical reasoning competency of clinical nursing instructors, final year Enrolled Nursing students and first year Enrolled Nursing students were -0.00, 0.09 and -0.02 respectively; with that the mean of clinical reasoning competency of the final year Enrolled Nursing students was highest. Table 5-11 showed that there was significant difference (p = 0.01) in clinical reasoning competency between the group of first year Enrolled Nursing students (M = -0.02, SD = 0.17) and the group of final year Enrolled Nursing students (M = 0.09, SD = 0.15). However, there was no significant difference (p = 0.13) found in the comparison between the clinical nursing instructors and the final year Enrolled Nursing students, and no significant difference was found when compared with the group of first year Enrolled Nursing students (p = 0.96). Nonetheless, according to Sawyer, S. F. (2009), the ANOVA tests can handle moderate violations of normality and equal variance by invoking the central limit theorem if there is a large enough sample size and a balanced design with equal sample sizes in each group, moreover, a retrospective power analysis is warranted after data is collected that aiming to determine the statistical power of the study which is based on the effect size and sample size that is particularly relevant for statistically non-significant findings, if the difference in sample sizes between the comparison groups is too large, this can affect the homogeneity of variance assumption as ANOVA is not as robust as when the difference is smaller. Thus, the result of non-significant difference may have been the result of inadequate statistical power or large difference in sample sizes. Hence, the result of non-significant difference between the group of clinical nursing instructors and the groups of Enrolled Nursing students might have been caused by the small sample size of the group of clinical instructors as well as the different samples sizes of the comparison groups. Therefore, in future study, using a larger sample of population and with equal sample sizes in each group for comparison is necessary.

6.5 Research Hypothesis: There is a difference in clinical reasoning competency between the final year Enrolled Nursing students and the first year Enrolled Nursing students.

Table 5-10 shows that a significant difference (p = 0.01) in clinical reasoning competency between the group of first year Enrolled Nursing students (M = -0.02, SD = 0.17) and the group of final year Enrolled Nursing students (M = 0.09, SD = 0.15) was found. It implies that the clinical reasoning competency of Enrolled Nursing students increases as the years of their clinical experience accumulated. Therefore, results reject the null hypothesis; rather to support the research hypothesis that more years of clinical experience can improve the clinical reasoning competency of the Enrolled Nursing students.

6.6 Research Hypothesis: There is a difference in clinical reasoning competency between the clinical nursing instructors and the Enrolled Nursing students.

Table 5-12 shows that there was no significant difference (p = 0.13) found in the comparison between the clinical nursing instructors and the final year Enrolled Nursing students, and similar result of non-significant difference was also found when compared with the group of first year Enrolled Nursing students (p = 0.96). Therefore, results failed to reject the null hypothesis and would not accept the alternative hypothesis. According to Sawyer, S. F. (2009), the ANOVA tests could handle moderate violations of normality and equal variance by invoking the central limit theorem if there was a large enough sample size and a balanced design with equal sample sizes in each group; if the difference in sample sizes between the comparison groups was too large, this could affect the homogeneity of variance assumption as ANOVA was not as robust as when the difference was smaller. Thus, the result of non-significant difference may be due to the large difference in two sample sizes (such as in this study, the comparison done between a group of 15 clinical nursing instructors and a group of 41 Enrolled Nursing students). Moreover, competence was the abilities individuals possess that enabling nurses to perform their duties to meet the required professional standard; however, these abilities might improve or diminish over time (Beidler, 2001; Tabari-Khomeiran & Parsa-Yekta, 2007). Furthermore, some sample subjects might try random guessing answers to the question items or they were careless or rush to choose answer due to limited time or the environmental factors; thus, the result may not reflect the real situation. For that, in future study, using equal sample sizes in each group for group comparison is necessary.

6.7 Implications of the CRCAT on Nursing Profession

6.7.1 Nursing Education

As aforementioned, the Hong Kong hospital-based general nursing training schools have adopted the written examination and clinical assessment as determinators to the eligibility of the final year nursing students towards the licensing registration in the Hong Kong Nursing Council (HKNC). The results of written examination and clinical assessment are actually the raw score calculation, which are supposed to indicate as well as evaluate nursing students' learning outcome both achievement and difficulties after instruction, they are summative assessments in nature; however, a score standing alone has no meaning and also same scores may have different meanings (Son, K. C., 2013). Moreover, the judgement of ability distance between two students with raw score is definitely test dependent. If the test developer has created many difficult items which student A can answer correctly but student B cannot, then the ability distance between them will be high; on the other hand, if the test developer created many easy items, then the ability distance between them will be very close to zero. In other words, the results of assessments with raw score calculation cannot reflect nursing students' ability or clinical reasoning competency. Hunter & Arthur (2016) commented that the current clinical performance assessment tools were unable to adequately appraise students' clinical reasoning competency during clinical placement. Moreover, Wang (2010) emphasized that the judgment of person level of ability or competency using raw score was test dependent. In other words, whether the results of the current assessments used in the hospital-based general nursing training schools under the Hospital Authority in Hong Kong could truly reflect the nursing students' clinical reasoning competency remains doubtful. In fact, in nursing education, several education models related to clinical reasoning such as the Outcome Present State Test, OPT (Figure 2-1) and Self-Regulated Learning, SRL (Figure 2-2) have already been used as teaching planning and methodology to develop clinical reasoning ability for the nursing students. Moreover, the clinical reasoning cycle (Figure 2-3), which is a model of thinking process consisting of the steps: consider, collect, process, identify, establish, act, evaluate and reflect, has also used as a teaching methodology especially for the subjects of health assessment and nursing procedures, etc. to improve nursing students' specific knowledge and to guide their clinical reasoning skills. The clinical reasoning competency assessment tool (CRCAT) provides a new assessment tool to help evaluate nursing students' knowledge of clinical reasoning and understand their own clinical reasoning capability, and also identify the areas of improvement in teaching methodology and clinical coaching strategies. Moreover, the results of this study shows that there is significant difference in clinical reasoning competency between the final year Enrolled nursing students and the first year Enrolled nursing students. It implies that the more the years of clinical experience the Enrolled nursing students accumulated, the higher level of clinical reasoning competency they will reach. Moreover, in Bloom's revised taxonomy, the six levels of cognitive process are ordered from simple to complex and from concrete to abstract. The classification is often referenced as a progressive climb to higher level of thinking with the highest level being "creating". The lowest level in the taxonomy deals with simple knowledge acquisition. At this level, students simply memorize, recall, list, repeat information or knowledge they have learned. The cognitive complexity grows at every level. At the highest level, students are able to put parts together to form a whole as well as to make judgments about the values of ideas or the complex clinical situation. In order words, the nursing students have to go through the different levels of cognitive process from the simple level of remembering and then to the level of remembering and then to the levels of applying and analyzing before they reach the level of evaluating which is the highest level of clinical reasoning competency in the CRCAT. Therefore, the curriculum design of nursing education programs especially the theory input and the clinical practicum are the determinative factors leading to the learning outcome of clinical reasoning competency.

6.7.2 Nursing Practice

Although there is no recognized assessment tool used to assess and monitor the clinical performance of nursing professionals in the hospitals of Hong Kong Hospital Authority, there are still some hospitals and clinical departments have annual written assessment using some complex nursing procedures, which are the high risk procedures easily causing incidents. The annual written assessment was designed to evaluate the clinical knowledge of nursing professionals, such as insertion of nasogastric tube, blood transfusion and administration of medication, etc.; however, the design of written assessment is direct question and answer approach without clinical context, the outcome performance is raw score calculation; thus, the results cannot truly reflect the clinical reasoning competency of nursing staff. The clinical reasoning competency assessment tool (CRCAT) provides script-based clinical context with the revised Bloom's cognitive process as the indicators of clinical reasoning competency. It can be used as a clinical assessment tool to evaluate the clinical reasoning competency of the clinical nursing staff and the nursing students to ensure their clinical practice is safe and up to standard.

6.7.3 Nursing Research

In this study, the Rach partial credit model analysis shows that the person reliability is acceptable at 0.64 and the person separation is 1.34 which is lower than the good separation index implying that if the same group of sample persons was given another parallel set of items that measuring the same construct, the replicability of person ordering cannot be as expected (Bond and Fox, 2007); in other words, the number of targeted items are not enough spread of ability among the persons, and more targeted items should be developed. Also, Cronbach suggested that a high value of alpha was desirable if the assessment tool was used to assign a score, which was interpretable, to a person (Taber, 2016). Therefore, further modification of the CRCAT is necessary before used in other sample targets, such as more script-based clinical scenarios should be developed using new collected clinical cases with uncertainty to increase the number of items in CRCAT. Apart from the Enrolled Nursing students, the Registered Nursing students and the nursing staff at different levels with different years of clinical experience can also be the target groups in future study using CRCAT with same script-based clinical scenarios in order to see if there are significant differences in clinical reasoning competency among these groups of nursing professionals. Further, the CRCAT can also be used in other countries outside Hong Kong after the modification of the script-based clinical scenarios done according to the clinical situations of the countries.

6.8 Limitations of the CRCA Tool

6.8.1 Representative and Generalization

The main purpose of development of CRCAT was used to assess the clinical reasoning competency of the final year Enrolled Nursing students who were in transition to enrolled nurses, the convenience sampling method was used in this study. The target subjects were the Enrolled Nursing students only from the program of two-year nursing training held in one of the hospital-based schools of nursing training in Hong Kong. Sample subjects were selected based on availability and willingness to take part, and useful results had been obtained, but the results were prone to significant bias, those who volunteered to take part might be different from those who chose not to or not being selected as sample subjects (volunteer bias), thus, the sample might not be representative of other characteristics, such as age, sex, clinical experience in different clinical context, caring different patients with different types of health problems, etc. Moreover, the total number of subjects was 97 that was less than 100, the sample size was comparatively small. In principle, the differences become smaller as the sample increases, less sample size unstable results delivered; large sample size essential and identical results provided. The Masters' Partial-Credit Model, in which each item defined its own scoring scale, 100 sample subject responses per each item might be too few, the minimum sample size was within the range from 108 to 243, 150 sample size was sufficient for most research purposes (Linarcre, J. M., 1994). Moreover, the research results could only be used if it had relevance to fields and people outside the contexts studied. Without generalization, there would be no evidence-based practice, and the best strategy for obtaining a representative sample of subjects was to use probability methods, such as random sampling (Polit, D. F. and Beck, C. T., 2010). Therefore, further study with larger sample size and extended scope of study field using probability sampling method is necessary to improving the representative and generalizability of the clinical reasoning competency assessment tool (CRCAT).

6.9 Conclusion

In conclusion, the results of the present study were discussed above, and the research questions and the hypothesis were also addressed based on the discussion made. The ICC results showed acceptable reliability of CRCAT, the content validity and understandability were established, and the construct validity was tested by examining the relationship between the CRCAT and the PSI using Pearson product-moment correlation coefficient as well as testing the model-data fit using Winsteps analysis. The Rasch Partial Credit Model analysis showed that the CRCAT had a good model-data fit but increased sample size and extended sample subjects were needed in order to strengthen the person reliability and the item polarity. Moreover, the results of the study had impacts on nursing education, nursing practice and nursing research, and the results also implied that the clinical reasoning competency will increase as the year of clinical experience accumulated. It certainly gave some insights as well as deep and profound thoughts and reflection to the nursing professionals especially those who were at the administration level in both clinical and academic nursing fields. Continuing professional training and monitoring of the clinical reasoning competency of the clinical nursing staff were the essential measures to ensure clinical safe practice. The conclusion of the thesis was presented in the following chapter.

CHAPTER SEVEN

CONCLUSION

The number of medical incidents in Hong Kong has been increasing in recent years, some of them were related to the nursing practice. These incidents, indeed, have made the nursing professionals have deep reflection and profound insights into the professional nursing training including curriculum of nursing education, teaching methodology, clinical coaching and the assessment tools used. Practically, having clinical reasoning competency to make timely decision to ensure patients' safety. This is crucial to all nurses especially those who are working in the frontline. In fact, the Hong Kong hospital-based schools of general nursing training has adopted various pedagogical methods in their training program to develop and equip the nursing students with the clinical reasoning ability during the years of training; however, the final written examination and final clinical assessment they use for the eligibility of registration in the Nursing Council of Hong Kong are still the traditional question-answer approach and procedural-oriented with raw scores calculation that are summative in nature and definitely test dependent (Soh, 2013), they cannot reflect the nursing students' clinical reasoning competency. As commented by Hunter and Arthur (2016) that the current clinical performance assessment tools were unable to adequately appraise nursing students' clinical reasoning ability during clinical placement.

The aim of this two-phased study was to develop and validate an assessment tool (CRCAT) for assessing the clinical reasoning competency of Enrolled Nursing students. In phase one, the clinical reasoning competency assessment tool (CRCAT) was developed using script-based clinical reasoning process design as suggested by Lubarsky, Dory, Duggan, Gagnon and Charlin (2013) as well as the Bloom's revised educational objectives taxonomy as the scoring method for examining the clinical reasoning competency levels from the first level of remembering to the fifth level of evaluating. The Rasch partial credit model analysis was used to examine the model-data fit through diagnosing data and cleaning data to yield objective scales for persons and items (Wang 2010). Concepts of four building blocks suggested by Wilson (2005) was applied as conceptual framework for the construction of CRCAT. The content validity and understandability of CRCAT were also established in this phase. In phase two, the cross-sectional design method was used for exploration of data from a convenient sample with three groups including the clinical nursing instructors, the final year Enrolled Nursing students and the first year Enrolled Nursing students. They have completed, at the same allowed duration of time, both the clinical reasoning competency assessment tool (CRCAT), and the problem solving inventory (PSI) which was used as a standard for comparison.

In response to the research questions and research hypotheses, the study results have been comprehensively discussed in previous Chapter Six. The results of Intra-Class Correlation

Coefficient (ICC) was 0.829 and the 95% confident interval of ICC estimate between 0.679 and 0.909, it indicates that the clinical reasoning competency assessment tool (CRCAT) has acceptable reliability. The face validity was not used because of its weakest form of measurement. The content validity and the understandability were established in study phase one. The construct validity was tested using the Pearson product-moment correlation coefficient to assess the relationship between the CRCAT and the PSI, the results showed positive correlation but not significant ($r = 0.07$, $p = 0.52$). Results of Winsteps analysis using Rasch partial credit model showed that the CRCAT has a good model-data fit; however, increased sample size and extended scope of samples were necessary in order to strengthen the person reliability and the item polarity.

Analysis of one-way variance (ANOVA) was conducted for the comparison between groups and within groups to evaluate the clinical reasoning competency among the groups of clinical nursing instructors and Enrolled Nursing students. Results confirmed that there was significant difference between groups ($p = 0.01$). Results of Tukey's HSD test showed that there was significant difference between the final year Enrolled Nursing students and first year Enrolled Nursing students, which rejected the null hypothesis.; however, no significant difference was found between clinical nursing instructors and final year Enrolled Nursing students, and between clinical nursing instructors and first year Enrolled Nursing students, which failed to reject the null hypothesis. These results implied that the more the years of experience the

Enrolled Nursing students has, the higher the level of clinical reasoning competency the Enrolled Nursing students will reach. In other words, the duration of clinical practicum is one of the essential parts of training for the nursing students to develop their clinical reasoning competency.

In conclusion, the study results have positive impacts on nursing administration, nursing education, nursing practice and nursing research. The clinical reasoning competency assessment tool (CRCAT) certainly provides new concepts of using scripts and Bloom's revised cognitive levels to develop a clinical assessment tool to examining the clinical reasoning competency of nursing professionals. These new concepts can also be applied in other countries, the scripts can be modified based on their own clinical situations. Small sample size and narrow scope of sampling were the limitations of this study. Therefore, in future study, a larger sample size and extended scope of target samples are necessary in order to polish and improve the representative and generalizability of the CRCAT.

REFERENCES

Anderson, L. W. & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

Andrew H.Liu, Zeiger, R., Sorkness, C., Mahr, T., Ostrom, N., Burgess, S., Rosenzweig, J. C., Manjunath, R. (2007). Development and cross-sectional validation of the children asthma control test. *Journal of Allergy and Immunology*, *119*(4), 817-825.

Armstrong, G. E., Spencer, T. S., & Lenburg, C. B. (2009). Using quality and safety education for nurses to enhance competency outcome performance assessment: A synergistic approach that promotes patient safety and quality outcomes. *Journal of Nursing Education*, *48*(12), 686-693.

Ashworth, P. D. and Saxton, J. (1990). On "Competence". *Journal of Further and Higher Education*, *14*, 3-25.

Atas, O. K. and Altunay, S. A. (2012). Optimal Sample Size Determination for the ANOVA Designs. *International Journal of Applied Mathematics and Statistics*, *25*(10), 127-134.

Baker, L. and Brown, L.A. (1984). *Metacognitive skills and reading. In P. D. Pearson (Ed). Handbook of reading research (pp.353-394)*. New York: Longman.

Bassendowski, S. & Petrucka, P. (2009). Perceptions of select registered nurses of the continuing competence program of the Saskatchewan registered nurses' association. *Journal of continuing Education in Nursing*, *40*(12), 553-559.

Benner, P. (1982). From Novice to Expert. *The American Journal of Nursing*, *82* (3), 402-407.

Beltran-Aguilar, E. D. & Beltran-Neira, R. J. (2004). Oral diseases and conditions throughout the lifespan. I. Diseases and conditions directly associated with tooth loss. *General Dentistry*, *52*(1), 21-27.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: fundamental measurement in the human sciences.* Mahwah, N.J.: Lawrence Erlbaum Associates Publishers

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: fundamental measurement in the human sciences (2nd Ed.)*. New York, N.J.: Routledge.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: fundamental measurement in the human sciences (3rd Ed.)*. New York, N.J.: Routledge.

Brach, C., Hall, K. K. & Fitall, E. (2019). Cultural Competence and Patient Safety. Retrieved from https://psnet.ahrq.gov/perspective/cultural-competence-and-patient-safety

Calder, B. J., Phillips, L. W. and Tybout, A. M. (2019). Designing Research for Application. *Journal of Consumer Research*, *8*(2), 197-207.

Carr S.M. (2004). A framework for understanding clinical reasoning in community nursing. *Journal of Clinical Nursing*, *13*, 850-857.

Cerit, K., Keskin, S. T. & Ekici, D. (2018). Development of instrument of bullying behaviors in nursing education based on structured equation modelling. *Asian Nursing Research*, *12*(4), 245-250.

Charlin, B., Boshuizen, H.P.A., Custers, E.J. and Feltovich, P.J. (2007). Scripts and Clinical reasoning. *Medical Education*, *41*, 1178-1184.

Charlin, B., Vleuten, C. V. D. (2004). Standardized Assessment of Reasoning in Contexts of Uncertainty: The Script Concordance Approach. *Evaluation & The Health Professions*, *27*(3), 304-319.

Charlin, B., Brailovsky, C., Roy, L., Goulet, F. and Vleuten, C. V. D. (2000). The Script Concordance Test: A Tool to Assess the Reflective Clinician. *Teaching and Learning in Medicine*, *12*(4), 189-195.

Chee, J. D. (2015). Pearson's Product-Moment Correlation: Sample Analysis. ResearchGate. Retrieved from https://www.researchgate.net/publication/277324930

Chien, C-W., Brown, T. & McDonald, R. (2011). Rasch analysis of the assessment of children's hand skills in children with and without disabilities. *Research in Developmental Disabilities*, *32*, 253-261.

Cortina, J. M. (1993). What Is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*, *78*(1), 98-104.

Cowan, D. T., Norman, I. J., and Coopamah, V. P. (2005). Competence in nursing practice: A controversial concept- A focused review of the literature. *Nurse Education Today*, *25*(5), 355–362.

Deschenes, M. F., Charlin, B., Gagnon, R. and Goudreau, J. (2011). Use of a Script Concordance Test to Assess Development of Clinical Reasoning in Nursing Students. *Journal of Nursing Education*, *50*(7), 381-387.

Duggan, P. and Charlin, B. (2012). Summative assessment of 5th year medical students' clinical reasoning by script concordance test: requirements and challenges. *BMC Medical Education*, *12*(29), 1-8.

Epstein, R. M. & E. M. Hundert (2002). Defining and Assessing Professional Competence. *Journal of American Medical Association*, *287*(2), 226-235.

Farra, S., Smith, S., French, D. and Gillespie, G. (2015). Development of an assessment instrument to evaluate performance of the skills of decontamination. *Nurse Education Today*, *35*, 1216-1022.

Field, W. E., Gallman, J. L., Nicholson, R. & Dreher, M. (1984). Clinical competencies of baccalaureate students. *The Journal of nursing education*, *23*(7), 284-293.

Fonteyn, M., and Ritter, B. (2008). Clinical Reasoning in Nursing, p236, In: Higgs, J., Jones, M.A., Loftus, S. and Christensen, N. (2008). *Clinical Reasoning in the Health Profession (3rd Ed)*. Oxford: Butterworth-Heinemann.

Fournier, J. P., Demeester, A. and Charlin, B. (2008). Script Concordance Tests: Guildelines for Construction. *BMC Medical Informatics and Decision Making*, *8*(10), 1-7.

Fuchs-Lacelle, S. and Hadjistavropoulos, T. (2004). Development and preliminary validation of the pain assessment checklist for seniors with limited ability to communicate (PACSLAC). *Pain Management Nursing*, *5*(1), 37-49.

Furst, E. J. (1981). Bloom's taxonomy of educational objectives for the cognitive domain: Philosophical and educational issues. *Review of Educational Research*, *51*(4), 441–453.

Gagnon, R., Charlin, B., Coletti, M., Sauve, E. and Vleuten, C. V. D. (2005). Assessment in the context of uncertainty: how many members are needed on the panel of reference of a script concordance test? *Medical Education*, *39*, 284-291.

Garside, J. R. & Nhemachena, J. Z. Z. (2012). A concept analysis of competence and its transition in nursing. *Nurse Education Today*, 1-5.

Girot, E. A. (1993). Assessment of Competence in Clinical Practice--A Review of the Literature. *Nurse Education Today*, *13*(2), 83-90.

Granello, D. H. (2001). Promoting Cognitive Complexity in Graduate Written Work: Using Bloom's Taxonomy as a Pedagogical Tool to Improve Literature Reviews. *Counselor Education & Supervision*, *40*, 292-307.

Guine R.P.F., Duarte J., Ferreira, M., Correia, P., Leal, M., Rumbak, I. & Straumite E. (2016). Knowledge about dietary fibres (KADF): development and validation of an evaluation instrument through structural equation modelling (SEM). *Public Health*, *138*, 108-118.

Hart, R. (2010). Winsteps Person & Item Bubble Chart. Retrieved from http://raschmodelaudit.blogspot.com/2010/12/winsteps-person-item-bubble-chart.html

Hazard Munro, B. (1997). *Statistical Methods for Health Care Research (3rd Ed)*. Philadelphia: Lippincott-Raven Publishers.

Hepper, P. P. & Petersen, C. H. (1982). The development and Implications of a Personal Problem-Solving Inventory. *Journal of Counseling Psychology*, *29*(01), 66-75.

Hepper, P. P. & Baker, C. E. (1997). Applications of the problem solving inventory. *Measurement and Evaluation in Counseling and Development*, *29*(4), 229-241.

Higgs, J., Jones, M. A., Loftus, S. and Christensen, N. (2008). *Clinical Reasoning in the Health Professions (3rd Ed)*. Philadelphia: Elsevier Butterworth Heinemann.

Hoffman, K., et al. (2010). The design and implementation of an Interactive Computerized Decision Support Framework (ICDSF) as a strategy to improve nursing students' clinical reasoning skills. *Nurse Education Today*, 1-8.

Holster, T. A. & Lake, J. (2014). How High Can They Jump: An Introduction to Rasch Measurement. Faculty of Literature International Faculty of Arts and Sciences Fukuoka Women's University, 78(02), 19-45.

Humbert, A. J., Johnson, M. T., Miech, E., Friedberg, F., Grackin, J. A. & Seidman, P. A. (2011). *Medical Teacher*, *33*, 472-477.

Hunter, S., & Arthur, C. (2016). Clinical Reasoning of Nursing Students on Clinical Placement: Clinical Educators' Perceptions. *Nurse Education in Practice*, *18*, 73-79.

Huang, Y-P. & Flores, L. Y. (2011). Exploring the Validity of the Problem-Solving Inventory with Mexican American High School Students. *Journal of Career Assessment*, *19*(04), 431-441.

Jones, P.W., Harding, G., Berry, P., Wiklund, I., Chen, W.H. and Leidy, N. K. (2009). Development and first validation of the COPD Assessment Test. *European Respiratory Journal*, *34*, 648–654.

Kassim, N. L. A., Ismail, N. Z., Mahmud, Z. and Zainol, M. S. (2010). Measuring Students' Understanding of Statistical concepts using Rasch Measurement. *International Journal of Innovation, Management and Technology*, *1*(1), 13-19.

Kautz, D.D., Kuiper, R.A., Pesut, D.J., Knight-Brown, P. and Daneker, D. (2005). Promoting clinical reasoning in undergraduate nursing students: application and evaluation of the outcome present state test (OPT) model of clinical reasoning. *International Journal of Nursing Education Scholarship*, *2*(1), 0-19.

Kevin, W.E. (2004). What every teacher needs to know about clinical reasoning. *Medical Education*, *39*, 98-106.

Kim, H.-Y. (2014). Analysis of variance (ANOVA) comparing means of more than two groups. *Restorative Dentistry & Endodontics*, *39*, 74-77.

Klein, C.J. (2006). Linking Competency-based assessment to successful clinical practice. *Education Innovation*, *45*(9), 379–383.

Kourmousi, N., Xythali, V., Theologitou, M. & Koutras, V. (2016). Validity and Reliability of the Problem Solving Inventory (PSI) in a nationwide Sample of Greek Educators. *Social Sciences*, *5*(25), 1-11.

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory into Practice*, *41*(4), 212-218.

Kuiper, R. A. and Pesut, D. J. (2004). Promoting cognitive and metacognitive reflective reasoning skills in nursing practice: self-regulated learning theory. Retrieved from https://doi.org/10.1046/j.1365-2648.2003.02921.x

Le Deist, F. D. & Winterton, J. (2005) What Is Competence? *Human Resource Development International*, *8*(1), 27-46.

Levett-Jones, T., et al. (2010). The 'five rights' of clinical reasoning: An educational model to enhance nursing students' ability to identify and mange clinical 'at risk' patients. *Nurse Education Today*, *30*, 515-520.

Linacre, J. M. (1994). Sample Size and Item Calibration Stability. *Rasch Measurement Transactions*, *7*(4), 328.

Lubarsky, S., Dory, V., Duggan, P., Gagnon, R. and Charlin B. (2013). Script concordance testing: From theory to practice: AMEE Guide No. 75. *Medical Teacher*, 1-10.

Lucas, J.W. (2003). Theory-Testing, Generalization, and the Problem of External Validity. *Sociology Theory*, *21*(3), 236-253.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*, 149-174.

Masters, G.N. (1988). *Measurement Models for Ordered Response Categories. In: Langeheine, R. & Rost, J. Latent Trait and Latent Class Models*. New York: Springer, Boston, MA.

Maydeu-Olivares, A. & D'Zurilla, T. J. (1997). The Factor Structure of the Problem Solving Inventory. *European Journal of Psychological Assessment*, *13*(03), 206-215.

Meterissian, S., Zabolotny, B., Gagnon, R., Charlin, B. (2007). Is the script concordance test a valid instrument for assessment of intraoperative decision-making skills? *The American Journal of Surgery*, *193*, 248-251.

McCrum-Gardner, E. (2008). Which is the correct statistical test to use? *British Journal of Oral and Maxillofacial Surgery*, *46*, 38-41.

Moghabghab, R., Tong, A., Hallaran, A. and Anderson, J. (2018). The Difference Between Competency and Competence: A Regulatory Perspective. *Journal of Nursing Regulation*, *9*(2), 54-59.

Munzenmaier, C. and Rubin, N. (2013). *Perspective Bloom's Taxonomy: What's Old is New Again. The elearning Guild in Research*. California: Scott Hanson.

Nkanginieme, K. E. O. (1997). Clinical Diagnosis as a Dynamic Cognitive Process: Application of Bloom's Taxonomy for Educational Objectives in the Cognitive Domain. Medical Education Online, 2(1), 106. Retrieved from http://www.utmb.edu/meo/

Nouh, T., Boutros, M., Reid, S., Pace, D., Walker, R., MacLean, A.,……Meterissian, S. H. (2012). The script concordance test as a measure of clinical reasoning: a national validation study. *The American Journal of Surgery*, *203*(4), 530-534.

Nota, L., Hepper, P. P., Soresi, S., & Hepper, M. J. (2009). Examining Cultural Validity of the Problem-Solving Inventory (PSI) in Italy. *Journal of Career Assessment*, *17*(04), 478-494.

Nurulhuda, S., Yasin, T. M., Fauzi, M., Yunus, M and Ismail, I. (2018). The Use of Rasch Measurement Model for the Validity and Reliability. *Journal of Counseling and Educational Technology*, *1*(2), 22-27.

Ostertagova, E. and Ostertag, O. (2013). Methodology and Application of One-way ANOVA. *American Journal of Mechanical Engineering*, *1*(07), 256-261.

Phillips, A. W., Smith, S. G. and Straus, C. M. (2013). Driving deeper learning by assessment: An adaptation of the revised Bloom's taxonomy for medical imaging in gross anatomy. *Academic Radiology*, *20*(6), 784-789.

Polit, D. F. and Beck, C. T. (2010). Generalization in Quantitative and qualitative Research: Myths and Strategies. *International Journal of Nursing Studies*, *47*, 1451-1458.

Portney, L. G. & Watkins, M. P. (2009). *Foundations of Clinical Research Application to Practice (3rd Ed)*. London: Pearson Education Inc.

Ramaekers, S., Kremer, W., Pilot, A., Beukelen, P. V. and Keulen, H. V. (2010). Assessment of competence in clinical reasoning and decision-making under uncertainty: the script concordance test method. *Assessment & Evaluation in Higher Education*, *35*(6), 661-673.

Sawyer, S. (2009). Analysis of Variance: The Fundamental Concepts. *The Journal of Manual & Manipulative Therapy*, *17*(02), 27-38.

Seddon, G. M. (1978). The properties of Bloom's Taxonomy of Educational Objectives for the cognitive domain. *Review of Educational Research*, *48*(2), 303–323.

Schober, P. (2018). Correlation and Coefficient: Appropriate Use and interpretation. Retrieved from https://www.researchgate.net/publication/323388613

Seaman, M. (2011). Bloom's Taxonomy: Its Evolution, Revision, and Use in the Field of Education. *Curriculum and Teaching Dialogue*, *13* (1/2), 29-43.

Siegle, C. B. H. & Cardoso de Sa, C. D. S. (2018). Concurrent validity between instruments of assessment of motor development in infants exposed to HIV. *Infant Behavior and Development*, *50*, 198-206.

Simmons, B. (2010). Clinical reasoning: concept analysis. *Journal of Advanced Nursing*, *66*(5), 1151-1158.

Soh, K. C. (2013). How much can we trust test scores? *Hong Kong Teachers' Centre Journal*, *12*, 131-148.

Su, W. M. & Osisek, P. J. (2011). The Revised Bloom's Taxonomy: implications for educating nurses. *Journal of continuing education in nursing*, *42*(7), 321-327.

Suciu, A. I. and Mata, L. (2011). Pedagogical Competences-The Key to Efficient Education. *International Online Journal of Education Sciences*, *3*(2), 411-423.

Sun, Y., Arning, M., Bochmann, F., Borger, J. and Heitmann, T. (2018). Development and validation of a practical instrument for injury prevention: the occupational safety and health monitoring and assessment tool (OSH-MAT). *Safety and Health at Work*, *9*, 140-143.

Taber, K. S. (2016). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. Science Education Centre, Faculty of Education, University of Cambridge, 6, 1-24.

Tabari-Khomeiran, R., Kiger, A., Parsa-Yekta, Z. & Ahmadi, F. (2007). Competence development among nurses: the process of constant interaction. *Journal of continuing Education in Nursing*, *38*(5), 211-218.

Teo, D. C. H., Suarez, L. & Oei, T. P. S. (2018). Psychometric properties of the problem solving inventory in a Singapore young male adult sample. Retrieved from https://doi.org/10.1007/s12144-018-0073-7.

Teodorescu, T. (2006). Competence versus competency: What is the difference? Retrieved from https://doi.org/10.1002/pfi.4930451027

Trevethan, R. (2016). Intraclass correlation coefficients: Clearing the air, extending some cautions, and making some requests. *Health Services and Outcomes Research Methodology*, *17*, 127-143.

Victor-Chmil, J. (2013). Critical Thinking versus Clinical Reasoning Versus Clinical Judgement: Differential Diagnosis. *Nurse Educator*, *38*(1), 34-36.

Waehrens, E. E. and Fisher, A. G. (2009). Developing linear ADL ability measures based on the ADL Taxonomy: A Rasch analysis. *Scandinavian Journal of Occupational Therapy*, *16*(3), 159-171.

Wang, W. C. (2010). *Recent Developments in Rasch Measurement*. HK: HKIEd.

Warden, V., Hurley, A. C. and Volicer, L. (2003). Development and psychometric evaluation of the pain assessment in advanced dementia (PAINAD) scale. *Journal of American Medical Directors Association*, *4*(1), 9-15.

Walton, D. (1990). What is Reasoning? What Is an Argument? *The Journal of Philosophy*, *87*(8), 399-419.

Watson, R., Stimpson, A., Topping, A. and Porock, D. (2002). Clinical Competence Assessment in Nursing: A Systematic Review of the Literature. *Advanced Nursing*, *39*(5), 421-431.

Wilson, M. (2005). *Constructing Measures An Item Response Modeling Approach*. NY: Taylor & Francis Group, LLC.

YasinA, R. M., Yunusa, F. A. N., Rusc, R. C., Ahmadd, A. and Rahimb, M. B. (2015). Validity and Reliability Learning Transfer Item Using Rasch Measurement Model. *Procedia-Social and Behavioral Sciences*, *204*, 212 – 217.

Clinical Reasoning Competency Assessment Tool (CRCAT): Assessment of clinical reasoning

Date of filling the CRCAT / /2016
(DD/MMM)

Please choose ***one most*** appropriate answer unless stated otherwise.
Mark your choice by **filling in the appropriate circle completely** as shown below**.**

	-2	-1	0	+1	+2
e.g.	◯	◯	◯	◯	◯

PART I (Clinical Scenarios):

1. A man aged 45 presents with burning abdominal pain was admitted for investigation.

If you are considering of	And then you find	Your consideration becomes				
Q1: peptic ulcer	Patient has chest pain	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: insertion of NG tube	Patient has dark stools	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: prop up patient	Patient has vomiting	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Almost unnecessary
-1	Not useful
0	Nor less nor more useful
+1	Useful
+2	Absolutely necessary

2. A 50 year old man presents with convulsion after chest X-ray taken.

If you are planning	And then you find	Your plan becomes				
Q1: to prepare IV therapy.	patient has generalized vigorously convulsion	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to do suctioning	patient looked pale	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to give oxygen	Patient's SpO2 > 90%	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

3.	**A 60 year old woman with history of liver cirrhosis was admitted due to hyperpyrexia of unknown origin (PUO), 2 liter of oxygen therapy given via nasal cannula. She was unexpectedly loss of consciousness while you were giving bedpan to her.**

If you are considering	**And then you find**	**Your hypothesis becomes**				
Q1: measuring vital signs for the patient	Patient was sweating	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: checking Haemoglucose level	Patient presented with four limbs convulsion	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: raising up bed-side rails	Patient looked pale	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

4.	**A young lady aged 24 with IDDM was admitted for DM control.**

If you are planning	**And then you find**	**Your plan becomes**				
Q1: to do haemoglucostix test	Patient's heart rate is 94 beats per minute.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to give 100 ml orange juice	The haemoglucostix value is 3 mmol/L	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to recheck haemoglucosestix test	Patient has severe headache	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

5. **A bedridden male patient diagnosed with Ca lungs was put on 100% oxygen therapy since the first day of admission, and he was oriented and independent. However, he was found unresponsive to both verbal and pain when you were trying to wake him up for oral medication this morning.**

If you were considering	And then you found	Your hypothesis became				
Q1: to assess GCS for this patient	The oxygen tubing was disconnected.	-2 ○	-1 ○	0 ○	+1 ○	+2 ○
Q2: to reconnect the oxygen tubing	Patient had drooling saliva	-2 ○	-1 ○	0 ○	+1 ○	+2 ○
Q3: to do oral nasal suctioning	Patient looked cyanotic	-2 ○	-1 ○	0 ○	+1 ○	+2 ○

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

6. **A 56 year old man presents with jaundice admitted for investigation.**

If you are thinking of	And then you find	Your hypothesis becomes				
Q1: hepatic jaundice	Patient's blood level of AST is high	-2 ○	-1 ○	0 ○	+1 ○	+2 ○
Q2: sickle cell anemia	Patient has abdominal pain and joint pains	-2 ○	-1 ○	0 ○	+1 ○	+2 ○
Q3: cirrhosis of liver	Patient's blood level of unconjugated bilirubin is high and conjugated bilirubin remains normal	-2 ○	-1 ○	0 ○	+1 ○	+2 ○

-2	Almost eliminated
-1	Less probable
0	The Information has no effect on the hypothesis
+1	More probable
+2	It can only be the hypothesis

7. **A 54 year old woman presents with chest pain and SOB admitted for investigation.**

If you are considering	And then you find	Your hypothesis becomes				
Q1: Angina Pectoris	ECG is normal	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: Ischemic heart disease (IHD)	Blood level of HDL cholesterol is high	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: Asthmatic attack	SpO2 > 94%	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Almost eliminated
-1	Less probable
0	The Information has no effect on the hypothesis
+1	More probable
+2	It can only be the hypothesis

8. **A large amount of fresh blood and blood clots was found on napkin when you were performing incontinence care for a female patient aged 80.**

If you are planning	And then you find	Your plan becomes				
Q1: to measure vital signs for the patient	More and more fresh blood was coming out from anus.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to lower down head of bed	Patient's blood pressure was 80/40 mmHg	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to give oxygen therapy according to doctor's prescription	Patient's SpO2 > 95%	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

9. **A female patient aged 72 transferred from OAH for OGD, oriented and responsive, now pending for results.**

If you are planning	And then you find	Your plan becomes				
Q1: to do mouth care for the patient	Patient has no natural teeth.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to sit out the patient	Patient has not been sitting out for one month.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to arrange the patient in 30 degree head up position	Patient was chesty	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

10. **An unconscious male patient aged 50 has received decompressive craniectomy and tracheostomy for one month.**

If you are planning	And then you find	Your plan becomes				
Q1: to do tracheostomy dressing	Patient has severe chesty coughing.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to do tracheostomy suctioning	Patient was on NG tube feeding.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to stop feeding	A large amount of yellowish secretion coming up from tracheostomy	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

11. **A young man aged 25 suffering from completed right-sided pneumothorax, and had pleurodesis performed with chest tube inserted on right side.**

If you are considering	And then you find	Your consideration becomes				
Q1: to check if the chest tube patent	No tidaling occurred in the water-seal chamber.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to assess the drainage in the tubing and chamber	Bubbling in the water-seal chamber	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to prop up patient	A small wedge is placed under patient's right shoulder blades	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

12. **A 55 year old man has just undergone transurethral resection of prostate (TURP) due to BPH, now was on continuing bladder irrigation.**

If you are planning	And then you find	Your plan becomes				
Q1: to measure vital signs	Patient has suprapubic discomfort	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to change a new NS irrigation bag for 2nd cycle	blood clots in urine drainage bag	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to prop up patient in Fowler's position	Urine output was limited	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

13.	**An 84 year old woman, conscious and alert, was transferred from OAH, and was found to have a big pressure ulcer on sacral area.**

If you are considering	**And then you find**	**Your consideration becomes**				
Q1: a stage III pressure ulcer	Subcutaneous fats at the right side bottom of the pressure ulcer	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to use hydrocolloid as dressing	It is a heavily draining wound	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to turn patient every two hours	Patient is resting on a low air-loss mattress	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

14.	**A 90 year old woman slipped down at home with head landed on the floor admitted for neurological observation and investigation.**

If you are thinking	**And then you find**	**Your intervention becomes**				
Q1: assessing patient's neurological condition	Patient is sleeping	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: arranging patient in semi Fowler position	Patient has high blood pressure and decreased pulse rate	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: arranging patient in lateral recumbent position	Patient has headache and vomiting	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

15. **A 10 year boy received tonsillectomy yesterday, and now presents low grade fever.**

If you are thinking	And then you find	Your intervention becomes				
Q1: to give 240 ml of apple juice to the child	No record of intake for 3 hours	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to give ice cream to the child	Child complains of ear pain	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to put child in recovery position	Child is splitting of blood	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

16. **A 48 year woman with diagnosis of type II DM and admitted for abdominal gap wound care.**

If you are planning	And then you find	Your plan becomes				
Q1: to give metformin 500mg to the patient before lunch as prescribed.	Patient's haemoglucostix test result was 5.9 mmol/L	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to clean the gap wound using chlorhexidine as prescribed.	The gap wound was in red color.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to use normal saline instead.	The gap wound had serosanguineous exudate	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

17. **A woman, 46 years old, suffering from Ca breast just returned to your surgical ward after mastectomy.**

If you are planning	And then you find	Your plan becomes				
Q1: to arrange patient in semi-fowler's position	Patient presented with SOB	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to support patient's neck in neutral position with small pillows on both sides	Patient's respiration rate was 26/min	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to give 24% oxygen via venturi mask	Patient's SpO2 = 90%	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

18. **A 90 year old woman suffering from frequent epistaxis admitted to your ward for cauterization.**

If you are planning	And then you find	Your plan becomes				
Q1: to put on PPE before doing nursing assessment.	Patient has nose bleeding again.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to apply direct external pressure to the patient's nares	Patient has severe sweating.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to provide tissues and an emesis basin to patient	Patient has no vascular access	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

19. **A woman 45-year-old had cholecystectomy yesterday, complained of nausea, and wanted to vomit.**

If you are planning	And then you find	Your plan becomes				
Q1: to prop up the patient.	Patient had distended abdomen.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to insert NG tube	Patient had undergone a laparoscopic cholecystectomy.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to give anti-emetic medication as prescribed	Patient had less feeling of nausea and vomiting.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

20. **An 82 year old woman has undergone an operation of total left hip replacement yesterday due to severe fracture hip after the accidental fall on street.**

If you are planning	And then you find	Your plan becomes				
Q1: to sit out the patient.	Patient is on IV therapy with a drain on left side and a Foley catheter.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q2: to teach patient to do ankle pumps exercise	Patient's lower limbs are swelling.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯
Q3: to give pain relief medication	Patient had nausea early in the morning.	-2 ◯	-1 ◯	0 ◯	+1 ◯	+2 ◯

-2	Strongly contra-indicated
-1	Not indicated
0	Nor less nor more indicated
+1	indicated
+2	Strongly indicated

CHRISTEE

PART II (basic demographic information):
Finally, please answer the following questions about your practice and yourself.

1.	**Age**		years		
2.	**Gender**	◯	Male	◯	Female

3. Education level

◯	F. 5	◯	F. 6 or F. 7		
◯	Diploma	◯	Bachelor	◯	Master or above

Others health care related training, please specify: ____

____

4. Have you been working in health care industry before attending the EN training program?

◯ YES ◯ NO (please go to Q7)

5.	**How long have you been working in hospital?**	years	months
6.	**How long have you been working in hospital as an APN?**	years	months

7. How many clinical practicum have you completed in EN training?

◯ CP1 ◯ CP2 ◯ CP3 ◯ CP4 ◯ CP5 ◯ CP6

END

Thank you very much for your participation

Used by Investigator
Subject ID:_______________
Date of filling ____(D)____(M) 2016

The Problem Solving Inventory

Directions: People respond to personal problems in different ways. The statements on this inventory deal with how people react to personal difficulties and problems in their day-to-day life. The term "problems" refers to personal problems that everyone experiences at times, such as depression, inability to get along with friends, choosing a vocation, or deciding whether to get a divorce. Please respond to the items as honestly as possible so as to most accurately portray how you handle such personal problems. Your responses should reflect what you actually do to solve problems, not how you think you should solve them. When you read an item, ask yourself: Do I ever behave this way? Please answer every item.

Read each statement and indicate the extent to which you agree or disagree with that statement, using the scale provided. Mark your responses by circling the number inside the table below each statement.

1. When a solution to a problem has failed, I do not examine why it didn't work.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

2. When I am confronted with a complex problem, I don't take the time to develop a strategy for collecting information that will help define the nature of the problem.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

3. When my first efforts to solve a problem fail, I become uneasy about my ability to handle the situation.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

4. After I solve a problem, I do not analyze what went right and what went wrong.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

5. I am usually able to think of creative and effective alternatives to my problems.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

6. After following a course of action to solve a problem, I compare the actual outcome with the one I had anticipated.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

7. When I have a problem, I think of as many possible ways to handle it as I can until I can't come up with any more ideas.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

8. When confronted with a problem, I consistently examine my feelings to find out what is going on in a problem situation.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

9. When confused about a problem, I don't clarify vague ideas or feeling by thinking of them in concrete terms.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

10. I have the ability to solve most problems even though initially no solution is immediately apparent.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

11. Many of the problems I face are too complex for me to solve.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

12. When solving a problem, I make decisions that I am happy with later.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

13. When confronted with a problem, I tend to do the first thing that I can think of to solve it.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

14. Sometimes I do not stop and take time to deal with my problems, but just kind of muddle ahead.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

15. When considering solutions to a problem, I do not take the time to assess the potential success of each alternative.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

16. When confronted with a problem, I stop and think about it before deciding on a next step.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

17. I generally act on the first ideal that comes to mind in solving a problem.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

18. When making a decision, I compare alternatives and weigh the consequences of one against the other.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

19. When I make plans to solve a problem, I am almost certain that I can make them work.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

20. I try to predict the result of a particular course of action.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

21. When I try to think of possible solutions to a problem, I do not come up with very many alternatives.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

22. When trying to solve a problem, one strategy I often use is to think of past problems that have been similar.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

23. Given enough time and effort, I believe I can solve most problems that confront me.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

24. When faced with a novel situation, I have confidence that I can handle problems that may arise.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

25. Even though I work on a problem, sometimes I feel like I'm groping or wandering and not getting down to the real issue.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

26. I make snap judgements and later regret them.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

27. I trust my ability to solve new and difficult problems.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

28. I use a systematic method to compare alternatives and make decisions.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

29. When thinking of ways to handle a problem, I seldom combine ideas from various alternatives to arrive at a workable solution.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

30. When faced with a problem, I seldom assess the external forces that may be contributing to the problem.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

31. When confronted with a problem, I usually first survey the situation to determine the relevant information.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

32. There are times when I become so emotionally charged that I can no longer see the alternatives for solving a particular problem.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

33. After making a decision, the actual outcome is usually similar to what I had anticipated.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

34. When confronted with a problem, I am unsure of whether I can handle the situation.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

35. When I become aware of a problem, one of the first things I do is try to find out exactly what the problem is.

1	2	3	4	5	6
Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Strongly Disagree

END

Thank you very much for your participation