EFFECTS OF ABDUCTIVE REASONING ON HYPOTHESIS GENERATION

EFFECTS OF ABDUCTIVE REASONING TRAINING ON HYPOTHESIS GENERATION ABILITIES OF FIRST AND SECOND YEAR BACCALAUREATE NURSING STUDENTS

By NOEMAN AHMAD MIRZA, B.Sc.N

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Doctor of Philosophy

McMaster University DOCTOR OF PHILOSOPHY (2015)

Hamilton, Ontario (Nursing)

TITLE: Effects of Abductive Reasoning Training on Hypothesis Generation Abilities of First and Second Year Baccalaureate Nursing Students.

AUTHOR: Noeman Ahmad Mirza, RN, BScN (McMaster University)

SUPERVISOR: Noori Akhtar-Danesh, Ph.D.

NUMBER OF PAGES: xv, 224

Lay Abstract

This study explored the effects of a training program on hypothesis generation abilities of nursing students. The training program aimed to teach students how to think more broadly about care situations. Student's hypothesis generation abilities were measured through the use of three care scenarios, each of which was presented before, immediately after and one-week after the training program. Only first and second year nursing students were included in the study. About half of the students were provided with the training while the other half were provided with informal discussion about hypothesis generation. After one-week, it was discovered that students who received the training had improved significantly in their ability to generate broad hypotheses. These students also generated hypotheses that were more accurate than the other group of students who did not receive the training. Due to the training, students' abilities in discovering the important aspects of the care situation also improved.

Abstract

Background: There is much debate on the best way to educate students on how to generate hypotheses to enhance clinical reasoning in nursing education. To increase opportunities for nursing programs to promote the discovery of accurate and broad-level hypotheses, scholars recommend abductive reasoning which offers an alternative approach to hypothetico-deductive reasoning. **Purpose:** This study explored the effects of abductive reasoning training on hypothesis generation abilities (accuracy, expertise, breadth) of first and second year baccalaureate nursing students in a problem-based learning curriculum. **Methods:** A quasi-experiment with 64 participants (29 control, 35 experimental) was conducted. Based on their allocation, study participants either took part in abductive reasoning training or informal group discussion. Three different test questionnaires, each with a unique care scenario, were used to assess participants' hypothesis generation abilities at baseline, immediate post-test and one-week follow-up. Content validity for care scenarios and other study materials was obtained from content academic experts. **Results:** Compared to control participants, experimental participants showed significant improvements at follow-up on hypothesis accuracy ($p=0.05$), expertise ($p=0.006$), and breadth ($p=0.003$). While control participants' hypotheses displayed a superficial understanding of care situations, experimental participants' hypotheses reflected increased accuracy, expertise and breadth. **Conclusion:** This study shows that abductive reasoning, as a scaffolding teaching and learning strategy, can allow nursing students to discover underlying salient patterns in order to better understand and explain the complex realities of care situations. Educating nursing students in abductive reasoning could enable them to adapt existing competencies when trying to accurately and holistically understand newer complex care situations. This could lead to a more holistic, person-based approach to care which will allow nursing students to see various health-related issues as integrated rather than separate.

Keyword: Abductive reasoning, clinical reasoning, hypothesis generation, accuracy, expertise, breadth, problem-based learning, nursing education

Acknowledgements

I would like to take this opportunity to thank my thesis supervisor, Dr. Noori Akhtar-Danesh, from the depth of my heart for his ongoing guidance and support during my doctoral studies at the School of Nursing at McMaster University. Noori, your constant encouragement was truly empowering and contributed immensely to my growth as a scholar and researcher. I would also like to thank Professor Charlotte Noesgaard who has been by my side since the beginning of my graduate studies. Charlotte, I have never learned as much about nursing education as I have from you. I hope that I can transcend this light onto others as you have done. I would also like to thank Dr. Lynn Martin who has also contributed to my learning since the beginning of my graduate studies. Lynn, I have always appreciated your insight and detailed feedback on the technical aspects of my writing. You have helped me become a better scholarly writer.

I would also like to take this opportunity to thank three other key committee members who have made great contributions to my learning and development as a nurse researcher. I am thankful to Dr. Carolyn Byrne for agreeing to join my committee and bringing a wealth of expertise to the table. Carolyn, I admire your sense of humour and your ability to lighten up the studious dialogues during our committee meetings. I have appreciated your thorough feedback especially during my thesis writing journey. I would also like to thank Dr. Basanti Majumdar (ex-committee member) for her guidance when I felt I was lost. Basanti, thank you for keeping me on track. I will always cherish the wonderful life experiences which you shared with me. I would also like to thank Dr. Eric Staples (ex-committee member) for his contributions to my learning in both the undergraduate and graduate levels. Eric, thank you for encouraging me to pursue graduate studies and showing continuous support during the early years of this journey.

In addition to my committee members, I would also like to thank several faculty members at the School of Nursing at McMaster University who offered their valuable expertise to help me conduct my research. These include (in alphabetical order by first name):

- Prof. Eileen Hanna
- Prof. Iris Mujica
- Prof. Louela Manankil-Rankin
- Prof. Lynda Bentley Poole
- Prof. Ola Lunyk-Child
- Prof. Tracey Jewiss

Finally, I would like to thank my family and friends, especially my brother, for believing in me even in times when I did not believe in myself. I could not have done it without your patience and ongoing support and encouragement. Thank you for being there every step of the way!

Table of Contents

List of Appendices

List of Figures

List of Tables

List of Abbreviations

PBL-PBL	Person-based learning within problem-based learning
PES	Problem, etiology, signs and symptoms
N	Number of participants
ICC	Interclass correlation coefficient
CVI	Content validity index
I-CVI	Item content validity index
ID	Identification
SPSS	Statistical Package for the Social Sciences
ANOVA	Analysis of variance
ANCOVA	Analysis of covariance
SD	Standard deviation

Declaration of Academic Achievement

Noeman Ahmad Mirza was the principal investigator and primary author of all chapters included in this thesis. Noeman's responsibilities, as the primary author, consisted of: conception, study design, data collection, data analysis, interpretation of study findings, and drafting of the written thesis.

Noori Akhtar-Danesh, Charlotte Noesgaard, Lynn Martin and Carolyn Byrne were all involved in this study. While Noori Akhtar-Danesh, Charlotte Noesgaard and Lynn Martin provided guidance on the study design, interpretation of study findings and refining of the written thesis, Noori Akhtar-Danesh further provided guidance on data analysis. Carolyn Byrne provided guidance on both the interpretation of study findings and refining of the written thesis.

CHAPTER 1: INTRODUCTION

Statement of the Problem

Seminal scholarly works in medicine and philosophy describe a *hypothesis* as an educated guess, proposition, hunch, or a plausible idea or theory which aims to explain certain patterns between a set of observable phenomena (Elstein, Shulman & Sprafka, 1978; Fann, 1970). Similarly, other works on hypothesis and human performance define *hypothesis generation* as the process where possible hypotheses are created to explain certain data or information (Fisher, Gettys, Manning, Mehle & Baca, 1983). Based on various nursing models, hypothesis generation is an integral component of scientific reasoning utilized by nurses to explain data presented in a given care situation (Hamers, Abu-Saad & Halfens, 1994). Strong abilities in hypothesis generation allow nurses to correctly plan, intervene, and evaluate nursing care for their clients (Needleman & Buerhaus, 2003).

Benner and colleagues (Benner, 1982, 1984; Benner, Tanner & Chesla, 2009) have extensively discussed the differences between novice (new graduates) and expert nurses. From these works, it is generally understood that nurses' knowledge, skills, and reasoning abilities develop with experience over time. While the rich mix of skills and experience allows expert nurses to intuitively grasp and respond to care situations, new graduates face challenges when analyzing data, detecting underlying patterns, sorting relevant information, prioritizing care issues, and formulating accurate explanations (e.g., hypotheses or diagnoses) about the presenting care situation (Li & Kenward, 2006; Purling & King, 2012).

Employers expect nursing schools to prepare graduates who can accurately detect and synthesize client data. When new nursing graduates face difficulty in meeting this expectation, their ability to safely manage a client's healthcare problem becomes questionable. For this reason, researchers encourage nurse educators in both clinical and academic settings to use active learning strategies which emphasize the application of knowledge and focus on enhancing novice nurses' and nursing students' abilities in synthesizing data, prioritizing care needs, and using appropriate interventions to manage health care problems (del Bueno, 1994, 2005; Purling & King, 2012).

While all nursing schools have their own unique teaching-learning strategies to enhance novice nursing students' abilities in synthesizing and prioritizing prominent client data, McMaster University's School of Nursing uses the *Person-Based Learning within Problem-Based Learning Model* (Figure 1.1), which is referred to as the PBL-PBL model (McMaster University BScN Handbook, 2012). Traditionally based on Barrows and Tamblyn's (1980) problem-based learning approach where learning is problem-driven, the revised PBL-PBL model focuses on the whole person rather than just the problem. According to Barrows and Tamblyn, the thinking process of problem-based learning is built on hypothetico-deductive reasoning, the focus of which is on testing pre-existing hypotheses in order to best explain a care situation (Elstein et al., 1978).

The PBL-PBL model outlines a cyclical, reflective hypothetico-deductive process which allows nursing students to encounter a care situation and use prior knowledge to generate early hypotheses in the form of care-related issues and their biopsychosocial mechanisms. By seeking out new information about the care situation, the students test their preliminary hypotheses. This leads to the confirmation or rejection of hypotheses. Confirmed hypotheses are used to explain the presented care situation. Newer care-related issues that need to be addressed are then selected and the PBL-PBL process begins again.

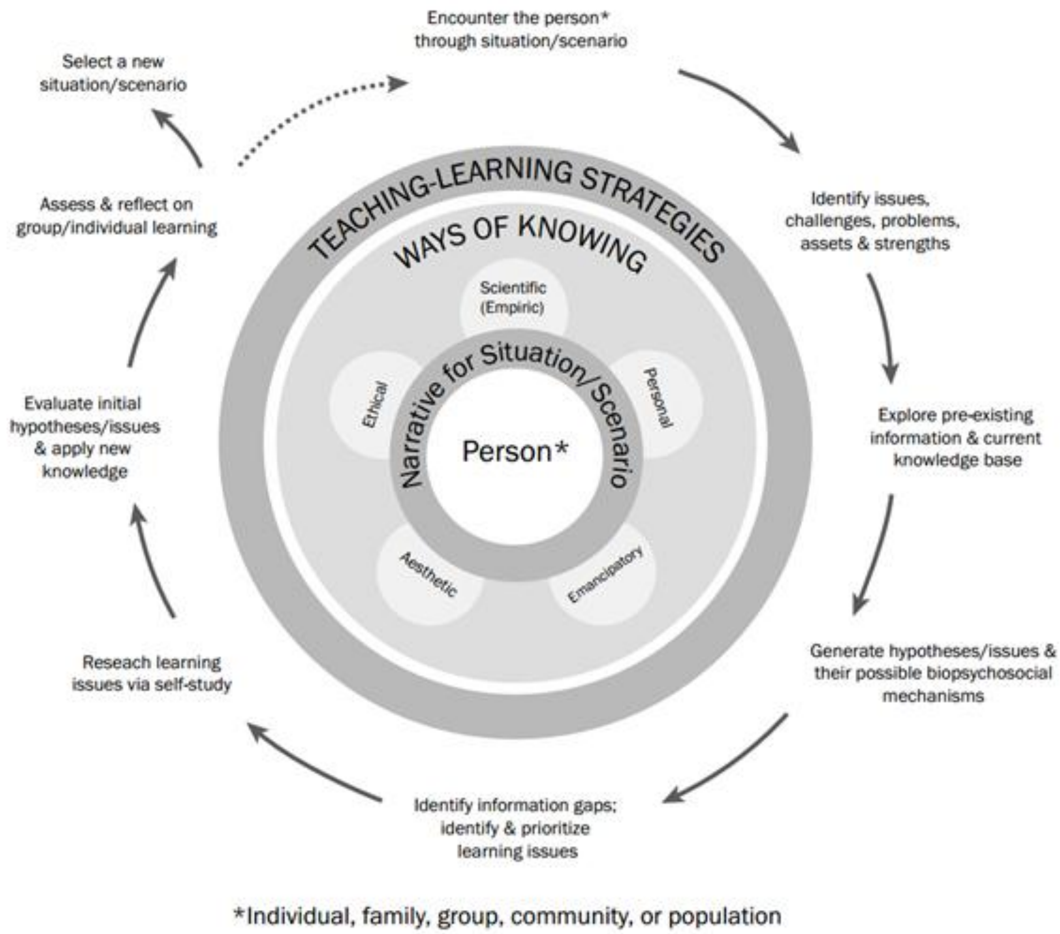


Figure 1.1. Person-based learning within a problem-based learning framework. Used with permission from the McMaster University BScN Handbook (2012).

Problem-based learning literature in nursing highlights the importance of testing previously known hypotheses and the evaluation of their accuracy, soundness or expertise, and breadth (biopsychosocial aspects) when resolving care situations (Ingram, Ray, Landeen, & Keane, 1998; Rideout, 2001a). However, no formal training on hypothesis generation is provided to nursing students at McMaster University. Instead, Socratic questioning (i.e., questions which probe deep-level thinking) is utilized to promote data analysis and hypothesis generation abilities among baccalaureate nursing students during class discussions within a problem-based learning tutorial (Jewiss, T., First Year BScN Chair, McMaster University, personal communication, 2012).

A clinician using hypothetico-deductive reasoning acquires a few initial cues from a care situation which help in the retrieval of a few (3 to 7) hypotheses from his or her memory. The clinician then tests these hypotheses against initial or newly acquired cues in order to confirm or refute them (Barrows & Tamblyn, 1980). According to critics (Fann, 1970; Haig, 1999; Vertue & Haig, 2008; Ward & Haig, 1997), hypothetico-deductive reasoning, upon which the thinking process of problem-based learning rests, focuses on the testing of predetermined hypotheses and does not focus on how hypotheses are created based on acquired cues that are specific to the client or person in the care situation. It is for this reason, these critics promote abductive reasoning, which is a unique form of reasoning involved in creating plausible explanations through the detection of underlying patterns (Eriksson & Lindström, 1997).

Abductive reasoning is a theoretical concept which could be useful in developing nursing knowledge (Raholm, 2010a, 2010b; Lipscomb, 2012). According to a concept analysis (Appendix 1) by Mirza, Akhtar-Danesh, Noesgaard, Martin and Staples (2014a), the practical implications of abductive reasoning have not been explored in nursing or in other clinical disciplines such as medicine and psychology. Due to the lack of guidelines to assist clinicians in applying abductive reasoning to care situations, Vertue and Haig (2008) developed a step-by-step approach to show how abductive reasoning could be applied to clinicians' practice within the discipline of psychology. Compared to the medical (Elstein et al., 1978) and nursing (Tanner, 2006) hypothetico-deductive approaches to clinical reasoning which focus on hypothesis testing, Vertue and Haig's abductive approach to clinical reasoning focuses on hypothesis creation. Table 1.1 illustrates the difference between the three different approaches.

Table 1.1: Abductive Reasoning Compared to other Reasoning Models

Abductive Reasoning (Vertue & Haig, 2008)	Hypothetico-Deductive Reasoning (Elstein et al.,1978; Barrows & Tamblyn, 1980)	Clinical Judgment (Tanner, 2006)
#1. Phenomena Detection		
a) Data Collection	#1. Cue Acquisition	#1. Noticing
b) Data Analysis	#2. Issue/Hypothesis Retrieval #3. Cue Interpretation	#2. Interpreting
#2. Inferring Causal Mechanisms	–	–
#3. Developing Causal Model	–	–
#4. Evaluating Causal Model	–	–
#5. Formulating the Case (Creation of final Hypothesis)	#4. Issue/Hypothesis Confirmation	Interpreting
–	–	#3. Responding
–	–	#4. Reflecting

As a hands-on approach to reasoning, Vertue and Haig's (2008) model is a theoretical idea that uses visual modeling techniques to allow the clinician to apply previous knowledge in order to analyze data, identify salient points, and formulate plausible hypotheses to explain care situations. However, this approach has neither been delivered in the form of training nor has it received research attention within the discipline of nursing. Since abductive reasoning focuses on creating hypotheses, it has been suggested that abductive reasoning could offer new ways of generating hypotheses beyond the traditional hypothetico-deductive approach to reasoning (Eriksson & Lindström, 1997; Lawson & Daniel, 2011; Raholm, 2010a; Rolfe, 1997; Vertue & Haig).

Nursing scholars consistently highlight the need for educators to use active learning strategies which emphasize students' abilities in: (1) detecting and interpreting important client cues, (2) applying previous knowledge, (3) synthesizing and prioritizing prominent or salient client data, (4) generating accurate, sound, and broad-scope hypotheses, and (5) choosing effective interventions to safely manage care situations (Benner, Sutphen, Leonard & Day, 2010; del Bueno, 1994, 2005; Purling & King, 2012). According to Mirza et al. (2014a) and Vertue and Haig (2008), abductive reasoning aims to assist clinicians in generating hypotheses by detecting, interpreting, and synthesizing important data. However, there is no research to support this claim within clinical disciplines, especially nursing.

Given that no formal training on hypothesis generation is provided to baccalaureate nursing students at McMaster University, training based on abductive reasoning could support hypothesis generation abilities of nursing students. Research exploration of the impact of this training on hypothesis generation abilities of baccalaureate nursing students can contribute to the limited literature on abductive reasoning in nursing and could shed new light on hypothesis generation within problem-based learning. Findings could suggest new directions for research and curriculum development in baccalaureate nursing education. Furthermore, if abductive reasoning training is able to strengthen nursing students' abilities in generating accurate, sound, and broad-scope hypotheses, it will also enhance their subsequent abilities in determining effective nursing interventions, which will allow them to safely manage care situations.

Conceptual Framework

When administering educational interventions to enhance performance (e.g., hypothesis generation abilities of baccalaureate nursing students), it is important to plan interventions in such a way that change in performance can easily be detected. Best practice guidelines on competence assessment from the College of Nurses of Nova Scotia outline several competency frameworks which currently exist in nursing and medical literature (Vandewater, 2004; Vess, 2007). These frameworks are summarized in Table 1.2. Some of these frameworks focus on practice-related competencies (e.g., assessment, intervention, critical thinking, leadership), while some focus on expertise-related competencies (e.g., novice versus expert nurses, nurses versus nurse practitioners). However, Miller's (1990) pyramid model for competence and performance assessment is notably superior in a classroom setting because of its focus on using knowledge to improve competence and practice. It further allows educators to assess students' reasoning within situations involving OSCE (objective structured clinical examinations), standardized client exercises, oral and written exams, and simulations (Vess, 2007).

Table 1.2: Competency Frameworks*

Reference	Name	Description
Lenberg, 1999	Lenburg's Competency Outcomes and Performance Assessment	• Advocates for curricula based on competency outcomes and performance assessment. • Eight practice competencies consist of: assessment and intervention, communication, critical thinking, teaching, human caring relationships, management, leadership, and knowledge integration skills.
Benner, 1984	Novice to Expert model	• Describes the level of proficiency expected from novice to expert practitioners. • Five-level process of skill acquisition and competence which consists of: novice, advanced beginner, competent, proficient, and expert.
Miller, 1990	Pyramid of Competence	• Illustrates key elements of clinical competence. • Four-level process of competence initiated by knowing of facts (knows), followed by the application of knowledge (knows how), then hands-on demonstration (shows how), and finally what the clinician does in practice.
Pew Health Professions Commission, 1998	Pew's Twenty-One Competencies for Future Clinicians	• Aims to assess competency in health care professionals. • Uses twenty-one competencies to examine values, skills, and knowledge of health professionals.
Canadian Nurses Association, 2006	National Framework for Continuing Competence for Registered Nurses	• Aims to guide the development of continuing competence programs that promote safe, ethical and competent care by registered nurses across Canada.
Ramirez, Tart & Malecha, 2006	Kane's Model-Based Practice Analysis and Test Specification	• Aims to obtain information and evaluate competence of emergency department nurse practitioner practice.

*Based on the work of Vandewater (2004) and Vess (2007).

Miller's (1990) model which emerged from the medical literature, attempts to show that competence predicts performance (Figure 1.2). The model has four levels: (1) *knows* (knowledge); (2) *knows how* (competence); (3) *shows how* (performance); and (4) *does* (action). Each of these stages is briefly described below:

1. *Knows (knowledge):* The base of Miller's model allows for the assessment of basic facts – i.e., what a student *knows* or the *knowledge* he or she possesses (Miller, 1990; van der Vleuten, 2000).
2. *Knows how (competence):* This stage is related to applied knowledge where the student is assessed on his or her level of *competence* – i.e., whether or not the student *knows how* to use that knowledge in a particular context or situation (Miller, 1990; van der Vleuten, 2000).
3. *Shows how (performance):* This stage focuses on hands-on functions where the student is assessed on his or her *performance*. In this stage, the student *shows how* the knowledge he or she gained can be acted out in a simulated situation (Miller, 1990; van der Vleuten, 2000).
4. *Does (action):* Following performance, the student is assessed on his or her *action* – i.e., what he or she *does* in actual clinical practice (Miller, 1990; van der Vleuten, 2000). In competence testing, this stage has been described as very challenging to measure due to limited reliable and valid instruments (Wass, Van der Vleuten, Shatzer & Jones, 2001).

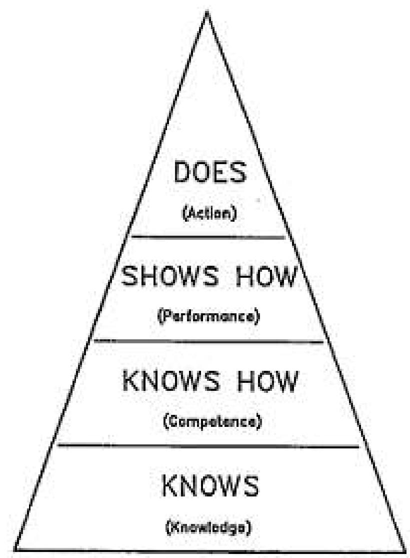


Figure 1.2. Miller's pyramid model of competence (known as the Framework for Clinical Assessment). This figure was taken from Miller (1990) with permission from Wolters Kluwer Health, Inc.

According to Miller (1990), healthcare graduates must be skilled in acquiring, analyzing, and interpreting data from multiple sources, and developing rational plans of care based on their findings. It is this functional adequacy – i.e., possessing sufficient knowledge, judgment, and skill to successfully complete a specific task – which Miller refers to as *competence*. Miller claims that several testing procedures require students to demonstrate the competence objective (i.e., the student *knows* and *knows how*), but sometimes ignore the performance objective (i.e., the student *shows how*) due to limited direct observations by instructors and because practice assessments emphasize the *product* (i.e., the correct diagnosis, the right intervention, etc.) rather than the *process* through which conclusions are arrived at (i.e., how hypotheses were generated, how diagnoses were formulated, and how plans of care were developed).

The focus of abductive reasoning is on the *process* of how conclusions, explanations, theories, hypotheses, or ideas are arrived at or how they come into existence (Haig, 2008; Lawson & Daniel, 2011; Lipscomb, 2012; Patel, Arocha & Zhang, 2005; Raholm, 2010a, 2010b). According to Miller (1990), the process component of performance needs more focus. Congruent with Miller's idea, abductive reasoning aims to engage clinicians in a process where they can use and apply their knowledge when collecting, analyzing, interpreting and synthesizing data from multiple sources; and detecting underlying patterns in order to generate explanations or hypotheses (Eriksson & Lindström, 1997; Raholm, 2010a; Vertue & Haig, 2008). As a result, abductive reasoning could be useful in determining the link between knowledge, competence, and performance of baccalaureate nursing students as they explore care situations.

Influential scholars examining the competence of new nursing graduates through their performance have concluded that although nursing graduates have passed competency examinations, their performance demonstrates below acceptable levels of competence in reasoning and judgment areas such as: analyzing data, detecting underlying causal patterns, determining relevant and salient issues, and formulating accurate and sound explanations (Benner et al., 2010; del Bueno, 1994, 2005; Li & Kenward, 2006; Purling & King, 2012). Similar to Miller's (1990) suggestion, even this pattern in nursing shows that although board examinations require students to demonstrate the competence objective (i.e., *knows* and *knows how*), the performance objective (i.e., *shows how*) is often not measured.

Van der Vleuten (2000) indicates that as skills become higher in Miller's (1990) pyramid, the assessments of these skills also become more clinically realistic. According to Wass et al. (2001), written and verbal examinations can be used to assess factual knowledge recall and applied knowledge. However, they suggest that more sophisticated methods involving cases and simulation are required to assess students' performance. This means that when Miller's pyramid of competence is being used as a guide to bring about change in performance (i.e., to enhance baccalaureate nursing students' abilities in hypothesis generation), it is important that educators utilize cases and simulation-based situations.

At McMaster University's School of Nursing, problem-based learning tutorials measure nursing students' *competence* through written and verbal assignments. Students' individual *performance* specific to their ability to generate accurate, sound and broad-scope hypotheses about care situations is assessed mainly through their professional practice courses. A student's tutorial performance in first and second year makes up 20-25% of the student's course grade and is measured based on three domains that include critical inquiry, group process, and professional conduct (McMaster University, 2013a, 2013b, 2013c, 2013d). However, these domains do not specify how and if nursing students are addressing performance objectives that are related to hypothesis generation within a problem-based learning tutorial where care situations are also explored.

Miller's (1990) pyramid model for competence and performance assessment is useful because it allows educators to observe whether change in performance (i.e., dependent variable) occurs as a result of an educational intervention at the competence stage (i.e., independent variable). Since testing procedures in problem-based learning tutorials predominantly measure competence of baccalaureate nursing students at the McMaster University School of Nursing, interventions aimed at improving hypothesis generation abilities through the use of abductive reasoning training must be measured at the performance stage. This can only be done by following the ideas of Wass et al. (2001) who suggest that controlled and standardized methods involving cases and simulation must be used to assess student performance.

When exploring abductive reasoning within Miller's (1990) model, baccalaureate nursing students' *knowledge* of hypotheses must first be assessed (i.e., does the student *know* what a hypothesis is?). Should students lack knowledge of hypotheses, education must be provided to warrant a baseline understanding of what a hypothesis is. To ensure students possess a sound understanding of hypotheses, their *competence* must then be assessed (i.e., does the student *know how* to generate a hypothesis?). This can be done by presenting students with brief care situations with prominent issues from which hypotheses can easily be generated. This is followed by the assessment of *performance* which determines whether students can detect, interpret, analyze, and synthesize important client data; and generate accurate, sound, and broad-scope hypotheses when they encounter care situations (i.e., can the student *show how* to generate a hypothesis?). This can be achieved through the use of care scenarios that require students to independently determine important issues and generate hypotheses to explain those issues.

The study reported within this dissertation used Miller's (1990) model in order to determine whether abductive reasoning training can improve hypothesis generation abilities of novice baccalaureate nursing students when compared to the standard hypothetico-deductive reasoning used within problem-based learning. Presented herein is a literature review that provides a historical overview of nursing education with an emphasis on how hypothetico-deductive reasoning is different from abductive reasoning. This is followed by a methodology section which describes the quantitative approach which was utilized to examine the effect of abductive reasoning on baccalaureate nursing students' hypothesis generation abilities. Finally, results and discussion are presented with implications and recommendations suitable to nursing education and clinical practice of nursing students.

CHAPTER 2: LITERATURE REVIEW

Historical Overview of Nursing Education

Professional nursing began in 1860 when Florence Nightingale founded the first school for nurses at St. Thomas's Hospital in London (Black, 2011). Before this, care of the sick was administered either by the client's relatives, by men and women in religious or military orders, or self-trained persons who held a low status in society (Chitty, 2011). The purpose of formal nursing education according to Nightingale (1859) was to enable nurses to master a unique body of knowledge before they engage in professional nursing practice. Since no professional nurses were available during the American civil war, untrained women responded to the appeal for nurses to care for the injured and dying soldiers. The need for professional nurses was realized and Nightingale's nursing school became the model of nursing education, which led to the emergence of formal nursing educational programs in America in 1873 (Black) and in Canada in 1874 (Elliot, Rutty & Villeneuve, 2013).

The last decade of the nineteenth century saw the emergence of nursing organizations such as the National League for Nursing, the American Nurses Association, and the International Council of Nurses (Black, 2011). By the end of the nineteenth century, there were hundreds of hospital-based nursing programs in America (Donahue, 1985), while a handful were established in Canada (Elliot et al., 2013). In the first decade of the twentieth century, licensure to protect the title of nurse and ensure standardization of nursing education programs were initiated in America in 1903 (Black) and in Canada with the formation of the Canadian Nurses Association in 1908 (Elliot et al.). While thousands of nurses joined the nursing workforce to meet the strong demand for nursing services posed by World War I, the influenza epidemic, the Great Depression, and World War II in the decades which followed, nurse leaders continued their efforts for the professionalization of nursing (Black).

Since nursing training was controlled by hospitals, medical administrators opposed nurses' aim to self-regulate and move education to university. This was due to paternalist ideas that included the assumption that as senior professionals, doctors must determine standards for nursing practice, must preserve the subordinate position of nurses, and remain in authority in the medical hierarchy (Kinnear, 1994). Despite opposition from hospital administrators, nurse leaders were successful in attaining power to self-regulate (i.e., gain control of their education, licensing, and the nursing discipline). By 1920, university-based nursing training programs were established at several universities which included Dalhousie, McGill, Toronto, Alberta, and British Columbia (Kinnear). However, the hospital-based apprenticeship model of instruction framed within biomedical thinking continued to be the principal method of educating nurses while university courses focused on administration or specialized fields such as public health (Coburn, 1974; Monti & Tingen, 1999).

After university-based nursing education was established, nurse leaders of the 1920s and 1930s began focusing on the transition of nursing from an apprenticeship towards an educational method through the development of nursing curricula which, in addition to medical knowledge, also included knowledge of the social sciences, pharmacology, and formal instructions on nursing procedures (Alligood, 2010).

Recognizing the over-emphasis placed on practical skills and the neglect of theoretical foundations upon which practical skills must be built, The Committee on Education of the National League of Nursing Education (1918) created standards of nursing education that promoted the use of lectures and demonstrations. However, these pedagogies involved passive learning strategies within the classroom since authority was in the hands of teachers (Scheckel, 2009). Even nursing research activities were limited in the early decades of the twentieth century (Burns & Grove, 2004).

Behaviourism

In the 1950s, nurse educators began using the Tyler model of curriculum instruction in order to develop nursing curricula which would prepare competent nurses (Bevis, 1988; Romyn, 2001). Tyler’s (1949, p.1) behaviour model of education posed four questions which he believed must be answered when developing curricula:

1. What educational purposes should the school seek to attain?
2. What educational experiences can be provided that are likely to attain these purposes?
3. How can these educational experiences be effectively organized?
4. How can we determine whether these purposes are being attained?

Tyler’s model was rooted in *behaviourism*, a learning theory which assumed that the teacher, as an authoritarian figure, was in control of designing the learning environment and shaping the student’s behaviour while the student was assumed to be a passive recipient with no control over the learning environment (Romyn, 2001). Behaviourism viewed learning as a change in or the acquisition of a behaviour (Pettigrew, 2015).

Behaviourism is known to have emerged from the works of three scientists, Thorndike, Pavlov, and Skinner, who examined learning as behaviour change through external conditioning and reinforcement (Schunk, 2012). Thorndike's (1911) popular research involved behavioural conditioning in which he placed animals in puzzle boxes and rewarded them with food if they figured out a way to escape the box. Repetition of this exercise enabled the animals to learn the behaviours which would enable them to escape the box. The animals were then able to apply this learning to different situations with similar setup.

Similarly, in his well-known experiment of *classical conditioning*, Pavlov (1927) introduced a dog with meat (unconditional stimulus) which caused salivation (unconditional response). While training the dog, Pavlov activated a metronome in the background (conditioned stimulus). Over time, the dog became conditioned to the metronome and would begin to salivate (conditioned response) upon its activation. Building on the work of Thorndike, Skinner (1953) proposed the idea of *operant conditioning* which stated that behaviour is followed by neutral, positive, or negative reinforcements. Positive reinforcement (reward) encourages certain behaviour and negative reinforcement (punishment) discourages certain behaviour; while neutral operant or response neither encourages nor discourages certain behaviours.

Behaviourist approaches to nursing education such as positive and negative reinforcements, behavioural objectives, lectures, programmed learning, and skill exercises were very influential in the 1950s. This was visible in the centralized lecture program that was initiated in Saskatchewan with the goal to provide nurses with four-month instructions in the basic sciences and to obtain recognition that nursing education belonged within the realm of education and not in the control of hospitals. Such initiatives led to increased movement of schools of nursing into the educational system, the closure of several hospital-based nursing programs in the 1960s, and the establishment of a national testing service from 1969 to 1970 to maintain national competency standards (Elliot et al., 2013).

Empiricism

In the 1950s, as nursing education in schools of higher education gained attention, higher education for nurses became a priority and nurses began to conduct research to develop a specialized body of knowledge unique to their discipline (Alligood, 2010). Influenced by the medical model, early nurse scientists adopted empiricism as the principal approach to developing nursing knowledge (Monti & Tingen, 1999). As a philosophy of science, *empiricism* stated that knowledge comes from sensory experience and evidence is discovered from experiments in which hypotheses and theories are tested against observations (Giuliano, 2003).

At the heart of the empiricist paradigm was the *scientific method* known as the *hypothetico-deductive method* of reasoning and knowledge development (Brody, 1993). According to Godfrey-Smith (2003, p.236), the hypothetico-deductive method of reasoning consisted of four steps: "(1) gather some observations, (2) formulate a hypothesis that would account for the observations, (3) deduce some new observational predictions from the hypothesis, and (4) see if those predictions are true. If they are true, go back to step 3. If they are false, regard the hypothesis as falsified and go back to step 2."

Cognitivism

Hypothetico-deductive reasoning originated from ideas on *information processing* which became popular in the 1950s when psychologists began shifting their focus from the traditional behaviourist approach to *cognitivism*, a learning theory which was influenced by Bruner's (1960, 1961) ideas on scaffolding. Cognitivism views learning as a change in mental structures rather than a change in behaviour and recognizes that learning involves mental processes through which a learner organizes and makes sense of the information being learned (Pettigrew, 2015). During this time (in the 1950s, specifically 1958), the nursing process, a modified version of the scientific method (i.e., hypothetico-deductive reasoning), was also developed (Koutoukidis, Stainton & Hughson, 2013; Marriner-Tomey & Alligood, 2006).

While empiricism enabled nurses to establish the scientific basis of their profession, this approach did not reflect the discipline's values and beliefs around holism, person-centered care, and understanding of the human experience. Subsequently, nursing theorists began suggesting that the discipline of nursing embrace pluralism through various patterns of knowing (Carper, 1978). As a result, the interpretative paradigm was also adopted because of its phenomenological and existential schools of thought which focused on the meaning of human experience, holism, individualism, autonomy, and self-determination rather than reductionism, objectivity, manipulation, prediction and control dictated by the medical (scientific) model (Harbison, 1991; Monti & Tinger, 1999). While nurse scientists were embarking on new ways of knowledge development, nurse educators were re-evaluating traditional teaching and learning approaches.

Constructivism

In the 1960s, the behaviourist approach to education was being criticized by researchers who were finding contradicting evidence to that discovered by the pioneers of behaviourism (Tomic, 1993). Freire (1970) described the traditional behaviourist approach to teaching as oppressive (i.e., authoritative teacher, passive learner). Nursing scholars also pointed to this in their critiques of the traditional pedagogy (Allen, 1990; Bevis & Murray, 1990; Chinn, 1989). Nurse educators began turning to *constructivism*, a learning theory which viewed learning as an active, contextualized process of knowledge construction through experience and reflection (Pettigrew, 2015). Under this perspective, the learner was viewed as a reflective and active individual who used personal experiences along with the teacher's guidance to discover coherent and organized knowledge rather than being a passive recipient of information (Mayer, 2004).

The focus of this new view on student-centered learning through discovery and its association with experiential education was influenced prominently by the work of Dewey, Bruner and Vygotsky. Dewey's (1938) concept of *progressive education* relied on purposeful learning which occurs when students participate and take interest in learning. Dewey further proposed the idea of *experiential education* which indicated that education must build on real life experiences. This not only encouraged curricula that were relevant to students' lives and their experiences, but also promoted *experiential learning* of practical life skills (Pettigrew, 2015).

Dewey's thinking was further transformed by Bruner who introduced ideas about the facilitation of learning by teachers. Bruner (1961) discouraged the instruction of knowledge and encouraged the facilitation of students' thinking and problem-solving abilities which could also be applied to other situations (Bruner, 1960). He promoted inquiry-based *discovery learning* where students as active learners, must construct their own knowledge by drawing on previous experience and organizing information in categories through a coding system (Bruner, 1961). Bruner's (1960, 1961) works also influenced the role of teachers whom he encouraged must facilitate and guide the learning process by *scaffolding* – i.e., providing students information and allowing them to discover the relationships among the bits of information.

Similar to Bruner, Vygotsky (1978) also stressed that learning has a social, interactive nature. He recognized that more knowledgeable individuals provide cognitive structuring or scaffolding to learners until they can perform without guidance. Vygotsky referred to this cognitive structuring as the *zone of proximal development* (p.86). Although constructive approaches discouraged passive learning and encouraged active learning, they created challenges for both students and teachers. Students were required to become more autonomous in the learning process, while teachers had to act as facilitators of discovery rather than knowledge-bearers, and were required to provide relevant frameworks to students upon which students could construct their knowledge and understanding (Chambers, Thiekötter & Chambers, 2013).

Self-Directed Learning

The introduction of ideas on adult education and self-directed learning aimed to meet some of these challenges. Knowles (1975) explained *self-directed learning* as the initiative taken by the learner in which he or she could identify learning needs, seek out learning resources, and evaluate learning outcomes either independently or with the guidance from the teacher. This learning process enabled the learner to become autonomous while the teacher acted as facilitator of learning. Building on Dewey's ideas of experiential learning, Kolb (1984) proposed an experiential learning model based on cognition, reflection, experience and the feelings and emotions associated with experience in order to take students' learning styles into consideration. This sequential model consisted of four stages:

1. *Concrete experience (feeling).* Learner is exposed to a concrete experience.
2. *Reflective observation (watching).* Learner observes and reflects on that experience.
3. *Abstract conceptualization (thinking).* Learner formulates abstract concepts
4. *Active experimentation (doing).* Learner tests new concepts in new situations.

Although new ideas on teaching and learning began emerging since the 1950s, several nurse educators continued using classroom pedagogies embedded in behaviourist approaches. It was not until the 1980s when the launch of the *curriculum revolution* by the National League for Nursing demanded nursing programs to reject behaviourist approaches to nursing education. The league encouraged the facilitation of student learning through the creation of new pedagogies aimed at preparing students who could participate as leaders in health care reform with values that recognized the diversity which existed in society (Tanner, 2007). This movement, along with evidence-informed teaching-learning practices, discouraged traditional passive learning strategies and encouraged strategies which would promote active student-centered learning such as cooperative learning, service learning, and problem-based learning (Scheckel, 2009).

Problem-Based Learning

To overcome the traditional behaviourist approaches to education, the School of Nursing at McMaster University adopted the active, student-centered, problem-based learning pedagogy in the mid-1970s after the university's medical school developed and implemented the pedagogy in 1969 (Rideout, 2001b). As an alternative approach to the traditional teacher-centered method of education, small-group problem-based learning is a student-centered pedagogy where the learning process is stimulated by a problem which learners wish to resolve (Rideout & Carpio, 2001). It is built on the framework by Barrows and Tamblyn (1980) who define it as "the learning that results from the process of working toward the understanding or resolution of a problem" (p.18). They also refer to problem-based learning as "discovery learning" (p. xi).

The idea of discovery learning is influenced by Bruner's (1961) ideas on enabling students to construct their own knowledge through previous experience. According to Rideout and Carpio (2001), problem-based learning is also influenced by the work of Dewey and Vygotsky. Dewey's (1938) ideas on experiential and purposeful learning promote a learning atmosphere where students are engaged in resolving hypothetical problems based on real life situations. Vygotsky's (1978) idea that learning has a social and interactive nature is practised through group discussions where the teacher acts as a facilitator of learning and provides guidance or scaffolding to the learners. According to Crooks, Lunyk-Child, Patterson, and LeGris (2001), learners within problem-based learning are expected to engage in self-directed and purposeful learning as recommended by Knowles (1975).

Hypothetico-Deductive Reasoning

In their problem-based learning framework, Barrows and Tamblyn (1980) promoted empiricism-based hypothetico-deductive reasoning as the principal approach to exploring and explaining care situations. To support their framework, they referred to the work of Elstein et al. (1978) who conducted a series of experiments which investigated how physicians solved problems. Elstein et al.'s work suggested that physicians should process information like scientists by using hypothetico-deductive reasoning. After the publication of their works, the hypothetico-deductive approach to medical reasoning not only became the principal method of diagnostic reasoning in medicine and medical education (Coderre, Mandin, Harasym & Fick, 2003), it also made its way into nursing literature where it was promoted as the ideal method for problem solving and decision-making (Hamers et al., 1994: Tanner, Padrick, Westfall & Putzier, 1987).

Barrows and Feltovich (1987), Elstein et al. (1978), and Elstein and Bordage (1988) described hypothetico-deductive reasoning within a clinical context in four steps:

1. *Cue acquisition.* Initial data are collected from a care situation which aid in the identification of issues.

2. *Hypothesis generation.* A few tentative hypotheses are retrieved from memory which guide further data collection.
3. *Cue interpretation.* Collected cues are used to determine whether a hypothesis should be accepted or refuted.
4. *Hypothesis evaluation.* Pros and cons for all alternative hypotheses are examined until one hypothesis is selected as the main diagnosis based on prevalent evidence.

Within the context of scientific reasoning, a *hypothesis* is described as an educated guess, proposition, hunch, or a plausible idea or theory which aims to explain certain patterns or relationships between a set of observable phenomena (Elstein et al., 1978; Fann, 1970). *Hypothesis generation* is generally understood as the process in which possible hypotheses are formulated to explain certain data or information (Fisher et al., 1983). There are two types of hypotheses – descriptive and relational. *Descriptive hypotheses* are concerned with the existence of one variable while *relational hypotheses* relate two variables together. Relational hypotheses are further divided into correlational and explanatory hypotheses. *Correlational hypotheses* indicate that two variables occur together in some manner but one may affect the other. *Explanatory (causal) hypotheses*, however, imply that the presence of, or a change in one variable causes an effect on the other variable (Sumathi & Saravanavel, 2008).

The four stages of hypothetico-deductive reasoning (i.e., cue acquisition, hypothesis generation, cue interpretation, hypothesis evaluation) are also emphasized in problem-based learning (Barrows & Tamblyn, 1980; Rideout & Carpio, 2001) and

McMaster University School of Nursing's approach to problem-based learning (McMaster University BScN Handbook, 2012). Not only is hypothetico-deductive reasoning promoted in problem-based learning curricula, it is also promoted in curricula which do not utilize problem-based learning (Wong & Chung, 2002). This is further confirmed in a literature review of various diagnostic reasoning frameworks utilized in nursing which indicates that most nursing reasoning models build on the four stages of hypothetico-deductive reasoning (Hamers et al., 1994).

Within a problem-based learning tutorial in nursing, the purpose of the four stages of hypothetico-deductive reasoning is to allow students to: (1) encounter a care scenario which will present a number of clinical cues which will trigger the retrieval of certain concepts from memory storage; (2) apply previous knowledge to the presenting care situation in the form of hypotheses; (3) seek out more information related to areas of interest so initial hypotheses can either be confirmed or refuted; and (4) select one or a few main hypotheses so appropriate actions can be determined in order to resolve the problem (McMaster University BScN Handbook, 2012; Rideout & Carpio, 2001). The end result of this reasoning process is to grasp the prominent hypothesis or hypotheses which best explain the presenting care situation.

Hypothetico-Deductive Reasoning in Nursing

It is well-documented in nursing literature that terms such as clinical reasoning, clinical judgment, clinical decision-making, problem-solving, critical thinking, and diagnostic reasoning are used interchangeably to describe hypothetico-deductive reasoning (Tanner, 2006; Thompson, 1999; Thompson & Dowding, 2002). In her concept analysis on clinical reasoning, Simmons (2010) not only confirms her belief by exploring extensive literature, she also describes that this way of thinking comprises of two aspects: a complex cognitive process that helps a clinician understand a care situation, and the clinician's response which often involves a choice, decision, or resolution.

Within the last five years, the School of Nursing at McMaster University has transformed the problem-based learning model into the *Person-Based Learning within Problem-Based Learning Model*, in order to shift the focus of learning from being problem-driven to being person-driven. In the revised approach, the person in the care situation is at the centre of learning and is examined from a holistic lens which takes into consideration biopsychosocial mechanisms (McMaster University BScN Handbook, 2012). Despite this change from problem to person in the revised version of the problem-based learning pedagogy, the hypothesis testing attribute of hypothetico-deductive reasoning remains the principal approach to exploring and explaining care situations which is similar to other nursing reasoning models outlined by Hamers et al. (1994).

Tanner's (2006) *Clinical Judgment Model* which aims to help learners think like a nurse, is also used by the School of Nursing at McMaster University. While the initial two phases of this four-phase model, *noticing* and *interpreting*, are concerned with the process involved in determining the hypothesis that explains a care situation; the latter two phases, *responding* and *reflecting*, are concerned with determining and evaluating the action which resolves the issues in the hypothesis. While the terminology is different, this model is similar to the nursing process – i.e., assessment (noticing), diagnosis (interpreting), intervention (action), and evaluation (reflection). Several nursing process proposals from the 1970s and 1980s also attempted to utilize varying terminologies to describe similar phenomena – e.g., *problem identification* instead of *diagnosis*, *implementation* instead of *intervention*, etc. (see Hamers et al., 1994).

While Tanner's (2006) model appears to be similar to the nursing process, it can also be considered similar to hypothetico-deductive reasoning since the nursing process itself is believed to be a modified version of the hypothetico-deductive, or scientific, method (Koutoukidis et al., 2013). Although Dowie (1988) proposed that the nursing process is a linear inductivist-model of reasoning (i.e., gather, sort and classify data before making a decision), there is general consensus that it follows hypothetico-deductive reasoning (Hamers et al., 1994; Koutoukidis et al.; McCarthy, 1981). These viewpoints illustrate that common reasoning models within the empiricist paradigm are based on hypothetico-deductive reasoning whether or not learning takes place within a problem-based learning context.

Novice and Expert Clinicians

In their exploration of hypothetico-deductive reasoning, Patel and Groen (1986) claimed that expert clinicians use reasoning from data to hypothesis and generate more accurate diagnoses; while novice clinicians use reasoning from hypothesis to data and generate less accurate diagnoses. Reasoning from *hypothesis to data* (i.e., *top down* or *backward* reasoning) is referred to as hypothetico-deductive reasoning; while reasoning from *data to hypothesis* (i.e., *bottom up* or *forward* reasoning) is referred to as *pattern recognition*, where known patterns from previous situations are recognized in the present situation (Elstein, 1994). Norman, Trott, Brooks, and Smith (1994) challenged this idea and suggested that both novice and expert clinicians use mixed reasoning strategies. However, Elstein summarized these opposing views and maintained that experts use more pattern-recognition while novices use more hypothesis testing strategies.

Similar to the ideas presented by scholars of medical education (Elstein, 1994; Patel & Groen, 1986), novice nurses also use hypothetico-deductive reasoning while expert nurses use pattern recognition (Tanner 2006). Pattern recognition in nursing is referred to as *intuition* (Benner, 1984; Tanner) and comes from the interpretive paradigm (Monti & Tingen, 1999). Introduced in the 1980s by Benner, intuition is described as understanding or knowing which lacks reason or rationale (Banning, 2008; Benner & Tanner, 1987). However, nurses who use intuition have faced criticism for having no rationale for their actions (Harbison, 1991). While it is believed that hypothetico-deductive reasoning can be taught (Rideout & Carpio, 2001), it is impossible to teach intuitive judgment since it develops, according to Benner's ideas, only with experience. As a result, nurse educators use the hypothetico-deductive model of reasoning to teach novice nurses the process involved in clinical reasoning (Harbison).

Challenges of Hypothetico-Deductive Reasoning

Hypothetico-deductive reasoning has, however, drawn criticism in relation to how hypotheses or diagnoses are generated. The first arguments involve the role of prior knowledge. Medical scholars Norman, Brooks, Colle, and Hatala (1999), and nursing scholars Simmons, Lanuza, Fonteyn, Hicks, and Holm (2003) point out that prior knowledge, which leads to the generation of early hypotheses, can result in a deliberate search for additional data to support one's hypothesis. This could lead the clinician to neglect other data which may also be important. Furthermore, since hypothetico-deductive reasoning is used for solving problems in the absence of relevant prior knowledge, there is no guarantee that diagnostic accuracy will be achieved due to the hypothesis testing nature of this method of reasoning (Patel et al., 2005).

The second set of arguments against hypothetico-deductive reasoning pertains to the generation of hypotheses. The initial phase of hypothetico-deductive reasoning involves the generation of early hypotheses from limited cues. This gives the false impression that the reasoning process begins with hypothesis generation rather than a complete detection of underlying patterns or relationships within the data (Ward & Haig, 1997). According to Buckingham and Adams (2000), initial hypotheses based on incomplete data may be incorrect and could lead to the generation of inaccurate final hypotheses or diagnoses.

Since hypotheses guide further action (i.e., clinical intervention or academic learning), it is important to generate accurate hypotheses which explain a broad range of issues in a given care situation (Ingram et al., 1998). However, novice nurses and nursing students have difficulty generating hypotheses which can comprehensively explain care situations. This is shown in the research conducted by del Bueno (1994, 2005) who examined the diagnostic abilities of over 30,000 new nursing graduates across several American hospitals. She found that only 35% of nursing graduates met entry-to-practice expectations in clinical reasoning and those who did not meet expectations were: (1) unable to provide accurate and specific hypotheses that explained the presenting care situation; (2) unable to differentiate problems requiring immediate intervention; and (3) provided inaccurate explanations for their actions.

Li and Kenward (2006) conducted a national survey of nursing education and practices of newly licensed nurses with 7,497 nurse graduates (76.5% registered nurses, 23.5% practical nurses). The researchers discovered that 18 to 20% of graduates felt they had inadequate educational preparation in nursing functions which included: (1) analyzing multiple types of data when making decisions; and (2) detecting underlying pathophysiological patterns of a client's health. According to Levett-Jones et al. (2010), several critical client incidents in hospitals involved poor clinical reasoning by new graduate nurses who were unable to properly detect and comprehend deteriorating care situations. Levett-Jones et al. attributed this trend to existing teaching and learning approaches which, according to them, are failing to facilitate the development of proper clinical reasoning abilities among nursing students.

In their literature review which explored new graduate nurses' preparedness for recognizing and responding to care situations, Purling and King (2012) highlighted that nursing graduates face challenges when reasoning through care situations. These challenges included difficulties when: (1) processing relevant information; (2) initiating medical intervention; and (3) evaluating outcomes. Purling and King, and del Bueno (1994, 2005) emphasize that employers expect nursing schools to prepare graduates who can accurately detect and synthesize data presented in a care situation; and that the inability to do so could mean that a client's healthcare problems will unlikely be managed safely. These scholars ascribe these trends to the limited use of active learning strategies which emphasize the application of knowledge and focus on enhancing novice nurses' and nursing students' abilities in synthesizing data to comprehend the situations, prioritizing care needs, and using appropriate interventions to manage the situation.

According to Mann (2012), nursing students' inability to think like a nurse (i.e., make appropriate choices and decisions) occurs mainly because: (1) nursing students lack opportunity to practice their reasoning skills outside of the clinical context; and (2) reasoning exercises are covered through instruction but students are not taught how to use reasoning strategies to explore care situations. While this may be the case in conventional nursing education, problem-based learning curriculum is different because it enables students to learn how to use reasoning strategies within the classroom setting as well through the use of care scenarios. Since intuition cannot be taught (Harbison, 1991), the reasoning strategies taught in problem-based learning tutorials are based on hypothetico-deductive reasoning (Barrows & Tamblyn, 1980).

After reviewing several studies, Rideout and Carpio (2001) indicated that in addition to increased satisfaction with the problem-based learning pedagogy, there is no significant difference between the knowledge and performance of students in a problem-based learning curriculum when compared to students in a conventional curriculum. While nursing education experts promote the use of hypothetico-deductive reasoning within all curricula (Harbison, 1991; Tanner, 2006; Tanner et al., 1987), both problem-based learning curricula and conventional curricula continue to use hypothetico-deductive reasoning as the principal method of reasoning, decision-making and problem-solving (Rideout & Carpio; Wong & Chung, 2002), despite its inadequacy in preparing new nursing graduates who can appropriately and competently manage care situations in the practice setting (del Bueno, 1994, 2005; Levett-Jones et al., 2010; Li & Kenward, 2006; Mann, 2012; Purling & King, 2012).

According to critics of hypothetico-deductive reasoning such as Fann (1970), Haig (1999), Vertue and Haig (2008), and Ward and Haig (1997), hypothetico-deductive reasoning focuses on the testing of predetermined hypotheses and not on how hypotheses are created. An example of this in nursing is NANDA's (North American Nursing Diagnosis Association) predetermined, standardized nursing diagnoses which began in the 1980s after the organization was established in 1982 (NANDA, 2014). Although nurses test and apply predetermined hypotheses to care situations, these hypotheses have, however, been found to be: (1) vague in their ability to explain care situations, (2) irrelevant to specific care needs of the client, and (3) inaccurate in depicting the client's main problems (del Bueno, 1994, 2005).

Abductive Reasoning

Since hypothetico-deductive reasoning has its limitations, critics of hypothetico-deductive reasoning (Fann, 1975; Haig, 1999; Vertue & Haig, 2008; Ward & Haig, 1997) point towards abductive reasoning as an alternative approach to hypothesis generation within the context of clinical reasoning. According to Patel et al. (2005), *abductive reasoning* is a unique form of reasoning that is involved in creating hypotheses. In nursing, abductive reasoning is described as a form of inference that focuses on the creation of plausible explanations through the detection of underlying patterns (Eriksson & Lindström, 1997).

Abductive reasoning originated from the work of Aristotle and was re-introduced in the present age by the philosopher of science, Charles S. Peirce (Eriksson and Lindström, 1997). Peirce (1903, 1931-1958) described abductive reasoning as the first stage of inquiry in which explanatory hypotheses are created to understand a complex reality, after which they must be clarified and tested. While hypothetico-deductive reasoning emerged from empirical thought and intuition from interpretive thought, abductive reasoning comes from *pragmatism*. Pragmatism is a unique way of thinking in the philosophy of science which was conceptualized by Peirce (1903) who described it as a doctrine within which ideas could only be considered meaningful if they had practical consequences. Problem-based learning pedagogy reveals several underlying ideas and procedures which are influenced by pragmatist thinking.

A major pragmatist philosopher and educational theorist, Dewey (1938), introduced pragmatist ideas into education by promoting purposeful and experiential learning and curricula that were relevant to practical life experiences. These ideas have also made their way into problem-based learning where real life examples are explored in learning tutorials (Barrows & Tamblyn, 1980; Rideout & Carpio, 2001). Bruner's (1960, 1961) ideas about scaffolding and Vygotsky's (1978) ideas about cognitive structuring and learning as a social, interactive process are also found in problem-based learning. Since the learning context within problem-based learning is affected by student and facilitator characteristics (motivation, social behaviour, etc.), the nature of the care situation and the quality of discussion, facilitators are being encouraged to become more practical (i.e., 'pragmatic enablers') in order to shift the present approach of problem-based learning towards pragmatism (Silén, 2004; Wilkie, 2004).

Abductive reasoning emerged from pragmatism and is presently a theoretical concept within nursing (Raholm, 2010a, 2010b; Lipscomb, 2012). However, its effect on the reasoning process involved in resolving care situations within healthcare disciplines such as nursing, medicine and psychology has not been explored. Due to the lack of guidelines to assist clinicians in applying abductive reasoning to care situations, Vertue and Haig (2008) developed a step-by-step approach to show how abductive reasoning could be applied to clinicians' practice within the discipline of psychology. Similar to Elstein et al.'s (1978) hypothetico-deductive approach to clinical reasoning which focuses on hypothesis testing, Vertue and Haig's (2008) abductive approach to clinical reasoning focuses on hypothesis creation. The step-by-step abductive reasoning approach proposed by Vertue and Haig involves:

1. *Phenomena detection.* Collect and analyzes data in order to detect issues/phenomena.
2. *Inference of causal mechanisms.* Infers possible causes of the detected phenomena.
3. *Development of a causal model.* Based on causes, develop a causal model to suggest how various causal mechanisms could be related to the detected phenomena.
4. *Evaluation of the causal model.* Ensures all links between causal mechanisms and detected phenomena are coherent and supported by data.

5. *Formulation of the case.* Generate a comprehensive and integrated plausible explanation (a hypothetical description combining several hypotheses) which explicates how various phenomena and causal mechanisms are related to one another.

While hypothetico-deductive reasoning limits the number of hypotheses (3 to 7) that can be considered by an individual (Barrows & Tamblyn, 1980), abductive reasoning has been linked to *connectionism* which allows multiple hypotheses to be considered simultaneously (Ajjanagadde, 1993; Johnson, Zhang, & Wang, 1997; Roth, 1996; Wang, Johnson, & Zhang, 2006). Although Vertue and Haig's (2008) theoretical ideas on abductive reasoning do not discuss connectionist ideas, they do offer an alternative approach to hypothetico-deductive reasoning which aims to consider a wide range of information by allowing clinicians to apply previous knowledge, identify salient points of a new care situation and create hypotheses that explain the situation. Since abductive reasoning focuses on creating hypotheses, it may offer new insight into hypothesis generation when compared to the hypothetico-deductive approach of hypothesis testing (Eriksson & Lindström, 1997; Lawson & Daniel, 2011; Raholm, 2010a; Reed, 1995; Rolfe, 1997; Vertue & Haig, 2008).

While the focus of reasoning used in the School of Nursing at McMaster University remains on hypothesis testing, there is less emphasis on the process involved in the generation of that hypothesis. Socratic questioning is used to promote deep level thinking among nursing students when they explore care situations (Jewiss, T., First Year BScN Chair McMaster University, personal communication, 2012). However, this strategy is discouraged in scientific inquiry by one of the originators of the scientific method, Popper (1945), who views Socratic questioning as vague and more suited for entertaining questions of a philosophical nature (e.g., What is caring?). Abductive reasoning, on the other hand, is based on pragmatism which views that although truth is unachievable, one must at least get close to it – i.e., the best explanation may not be entirely true but may be the closest conceptualization to the truth (Fann, 1970; Raholm, 2010a).

Abductive Reasoning and Science Education

In basic science education, abductive reasoning (hypothesis generation) training has enhanced brain activation patterns among students and consequently, has improved problem solving and hypothesis generation abilities. In their study on the impact of hypothesis generation skills on learning-related changes in brain activation, Kwon, Lee, Shin, and Jeong (2009) used 18 undergraduate science students and examined their brains using functional magnetic resonance imaging (fMRI) before and after the training which was 60 minutes of instruction per week for two months. Based on an ideal list of answers, intervention participants obtained significantly higher scores on their hypotheses than participants in the control group. Researchers also found that hypothesis generation resulted in higher-order inferential processing.

Another study by Kwon, Jeong, and Park (2006) examined 290 fifth graders' hypothesis generation abilities about pendulum motion. Children were asked to conceptualize the cause that creates a difference in speeds of pendulum motion. Face validity of the test was obtained from a panel of experts which included faculty and graduate students. The researchers paid particular attention to children's hypotheses and the use of prior understanding. They discovered that while participants possessed prior understanding of pendulum motion, a significant number of participants failed to apply their prior understanding when generating hypotheses on the study task. Findings also suggested that children engaged in abductive reasoning were more successful in generating hypotheses. As a result, researchers recommended that students be trained in abductive reasoning training which will develop their hypothesis generation abilities.

Reasoning Training

By examining several studies, Wolf, Gruppen, and Billie (1988) made the claim that medical students and physicians often fail to select optimal data from care situations which results in premature diagnoses. In order to provide clinicians with a useful cognitive strategy for guiding data gathering, these scholars conducted a quasi-experiment in which they tested the usefulness of a competing-hypothesis heuristic. Participants (n=200) completed three cases. While the intervention group completed the first case at baseline and the remaining two immediately after the intervention, the control group completed all cases in succession. Both groups were similar at baseline, while significant improvement was found in the post-intervention hypotheses of participants in the intervention group. The researchers suggested that training involving cognitive strategies to guide clinical reasoning and hypothesis generation are useful, can be learned, and should be included in professional training.

Recognizing that formal explicit teaching on exceptional clinical reasoning is rarely provided in medical education, Round (1999) carried out a non-randomized controlled observational study of 186 senior medical students. In this study, effects of a teaching intervention on clinical reasoning skills were measured using the diagnostic thinking inventory. Results indicated that the intervention group performed significantly better on the diagnostic thinking inventory than the control group. This showed that explicit teaching about clinical reasoning skills can enhance how healthcare students think about care situations. Furthermore, the study promoted the idea that reasoning skills can be learned and taught.

Perkins (1987) had discussed the importance of teaching students about thinking frames which can guide and support students' thought processes when problem-solving and can also improve intellectual competence. Perkins claimed that learners cannot infer thinking frames from the role-modeling of their mentors; rather, thinking frames need to be acquired through instruction. Examples of thinking frames which have been explored in nursing studies consist of the hypothetico-deductive method of reasoning, the PES (problem, etiology, signs and symptoms) guide, and Engel's (1977) biopsychosocial model for viewing holistic aspects of a client's health.

Tanner (1982) conducted an experiment to test an instructional intervention on the diagnosis process and its impact on nursing students' diagnoses. The intervention was based on the principles of hypothetico-deductive reasoning. Results indicated that instructions on hypothetico-deductive reasoning had no significant difference on nursing students' diagnoses. Florin, Ehrenber, and Ehnfors (2005) also tested the impact of PES training on the quality of nursing diagnoses. They defined quality by the structure and relevance of PES components within a relational diagnostic statement and developed their own scale by combining two previously existing, psychometrically-sound scales. Florin et al. discovered that while no difference was found in the control group, the quality of diagnostic statements improved significantly in the experimental group. This showed the need to train nurses in diagnostic reasoning in order to enhance the accuracy of their documentation.

Hypothesis Generation Abilities

Ingram et al. (1998) conducted an intervention study to evaluate the effectiveness of a board game in helping baccalaureate nursing students improve their hypothesis generation abilities (i.e., accuracy, expertise, and breadth). In addition to the biopsychosocial thinking frame, Ingram et al.'s study also examined developmental, spiritual/cultural, and political/economic categories. No significant changes in hypothesis generation abilities were detected between control and intervention participants. The study outcomes were described as such:

1. *Accuracy.* The number of correct responses per care scenario (based on expert list of responses) as a percentage of the number of responses generated for that scenario.
2. *Expertise.* The number of correct responses per care scenario (based on expert list of responses) as a percentage of all ideal responses verified by experts for that scenario.
3. *Breadth.* The number of biopsychosocial areas which participants' responses covered per care scenario.

These outcomes have also been discussed by Rideout (2001a) and Elstein et al. (1978) in relation to hypothesis generation. However, these researchers have used alternative terms to describe these outcomes. Rideout describes accuracy as precision while Elstein et al. describe it as correctness and use the term efficiency to describe the definition of accuracy presented by Ingram et al. (1998). Breadth is viewed by Rideout as the biopsychosocial features of a care situation while Elstein et al. do not discuss breadth due to their focus on the narrowing down of data in order to come up with a specific medical diagnosis. Rideout refers to expertise as *appropriateness* related to the main features of a care scenario. This is supported by Margetson (1997) who defines expertise as "an ability to make *sound judgments* as to what is problematic about a situation, to identify the most important problems, and to know how to go about solving or at least ameliorating them" (p.38). However, Elstein et al. refer to Ingram et al.'s definition of expertise as *thoroughness*, which they concluded has little influence on accuracy.

Since issue and hypothesis generation have been used interchangeably in nursing literature (Ingram et al., 1998; McMaster University BScN Handbook, 2012), it is important to differentiate issue generation from hypothesis generation. With guidance from Barrows and Tamblyn (1980), Fisher et al. (1983), Rideout and Carpio (2001), and Sumathi and Saravanavel (2008), *issue identification* can be described as the formulation of a list of problems based on the perception and interpretation of information presented in a situation (i.e., descriptive hypotheses, one variable); and *hypothesis generation* can be defined as the automatic and creative process of using knowledge and experience to formulate provisional propositions which explain and relate specific issues (i.e., explanatory hypotheses, two or more variables).

Retention of Reasoning Training

While studies express the need to enhance nursing students' abilities in hypothesis generation, varying lengths of interventions ranging from weeks to days are mentioned. However, some studies indicate that short one-time intensive training sessions are sufficient for enhancing students' cognitive abilities in cognitive tasks. Ruiz-Primo (2000) outlines several studies conducted by her and colleagues to evaluate an intensive 50-minute concept mapping training program. She discovered that this intensive training program was effective in allowing over 300 high school science students and several science teachers to grasp concept mapping as a cognitive strategy for organizing their knowledge.

In another study on concept mapping training for microbiology students, Kinchin, De-Leij, and Hay (2005) found that a one-hour training session was sufficient to teach university students about concept mapping after which students were given 15 minutes to make concept maps for selected lectures throughout the term. Hay and Kinchin (2008) further claim that the concept mapping method can be taught in 20 minutes, after which another 30 to 40 minutes must be allocated to allow students to make satisfactory maps related to a given topic. However, long-term retention of cognitive abilities related to thinking frames (i.e., hypothetico-deductive reasoning, PES, biopsychosocial model, concept mapping) were not assessed by researchers.

As nursing scholars have highlighted, there is a need for educators to use active learning strategies which emphasize students' abilities in: (1) detecting and interpreting important client cues, (2) applying previous knowledge, (3) synthesizing and prioritizing prominent or salient client data, (4) generating accurate, sound (appropriate, competent, expert), and broad-scope hypotheses, and (5) choosing effective interventions to safely manage care situations (Benner et al., 2010; del Bueno, 1994, 2005; Ingram et al., 1998; Purling & King, 2012; Rideout, 2001a). While research in science education shows promise that abductive reasoning allows students to apply previous understanding, Vertue and Haig's (2008) thinking frame promotes the possibility of improved abilities in detecting, interpreting, and synthesizing client data upon which accurate, sound, and broad-scope hypotheses can be built.

The research study reported within this dissertation explored whether abductive reasoning can improve hypothesis generation abilities of novice baccalaureate nursing students enrolled in a nursing program where hypothetico-deductive reasoning is used as a principal approach to exploring care situations. This study aimed to address the question:

> *Does abductive reasoning training improve hypothesis generation abilities (accuracy, expertise, breadth) of first and second year baccalaureate nursing students in the Basic BScN Program at the McMaster University School of Nursing?*

In the PICOT (population, intervention, comparison, outcome, and time) format, this question can be stated as such:

Population. First and second year baccalaureate nursing students from the four-year basic BScN program at McMaster University.

Intervention. Abductive reasoning training based on Vertue and Haig (2008)'s framework.

Comparison. Standard group discussion on hypothesis generation.

Outcome. Accuracy, expertise, and breadth of hypotheses.

Time. 15-minute tests (pre-test, post-test, follow-up) and one-week follow-up.

Study objectives consisted of:

1. Examining the effect of abductive reasoning training on the hypothesis generation abilities (accuracy, expertise, breadth) of first and second year baccalaureate nursing students in the Basic BScN Program at the McMaster University School of Nursing.
2. Comparing hypothesis generation abilities (accuracy, expertise, breadth) of first and second year baccalaureate nursing students in the Basic BScN Program at the McMaster University School of Nursing.

The main study hypotheses that were tested in the study were:

1. Abductive reasoning training will improve participants' hypothesis accuracy.
2. Abductive reasoning training will improve participants' hypothesis expertise.
3. Abductive reasoning training will improve participants' hypothesis breadth.

In order to explore these questions, a quantitative methodology was developed.

CHAPTER 3: METHODOLOGY

This study aimed to determine whether abductive reasoning can improve hypothesis generation abilities of novice baccalaureate nursing students when compared to the standard hypothetico-deductive reasoning. It also sought out to examine any differences between first and second year baccalaureate nursing students which could have resulted from abductive reasoning. While the CONSORT (Consolidated Standards of Reporting Trials) statement is recommended for reporting randomized controlled trials (Begg et al., 1996), the TREND (Transparent Reporting of Evaluations with Non-randomized Designs) statement is recommended for reporting non-randomized intervention studies (Des Jarlais, Lyles, Crepaz & TREND Group, 2004). In their comparative analysis (Appendix 2), Mirza, Akhtar-Danesh, Staples, Martin and Noesgaard (2014b) indicate that the TREND statement could be used as a guideline when reporting studies that use educational interventions. Therefore, the TREND statement was consulted when reporting the present research study.

Pilot Study

The pilot study was conducted in the fall (October and November) of 2013 with the aim to examine the practicality of study materials, to make any changes to study procedures, and to estimate a preliminary sample size. To conduct the pilot study, a simple one-time pre-post experiment was conducted with 20 participants who completed all study materials and received the intervention in one sitting. The length of each study session was approximately 2 hours. Participants were able to comprehend the intervention and certain teaching parts of the intervention were simplified to assist with understanding.

Participants verbalized that the new approach to exploring care situations (i.e., abductive reasoning) helped them integrate previous knowledge when trying to generate hypotheses for the presenting care situations. No follow-up was done because the focus of the pilot study was to determine the practicality of study materials, the intervention, and to note possible challenges with student recruitment.

Study Design

A non-randomized, quasi-experimental design with a control group and a follow-up was employed for the study in Winter 2014 (Figure 3.1).

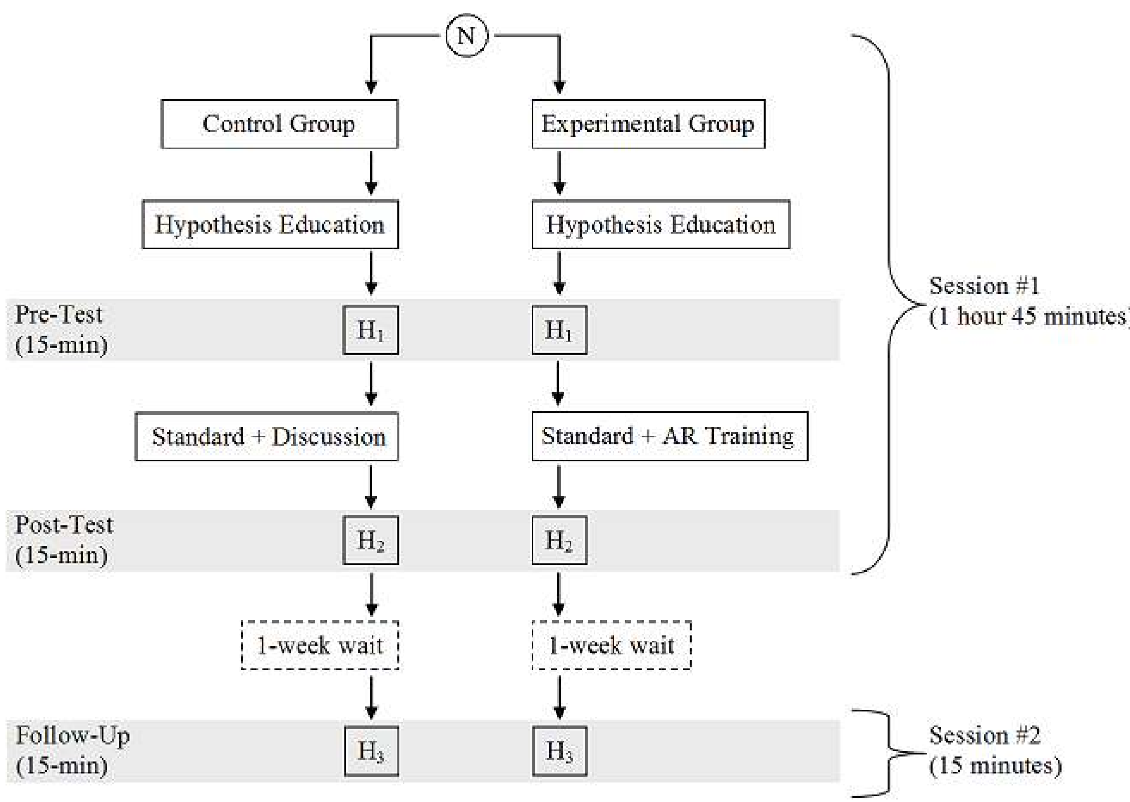


Figure 3.1. Schematic diagram of study design; where N = number of participants, H = test questionnaires, and AR = abductive reasoning.

Randomized controlled trials are not feasible in educational settings due to the risk of contamination and practical constraints related to student recruitment and their commitment to participation in research studies (de Anda, 2007; Markert, O'Neill, & Bhatia, 2003). A control group in addition to the intervention group from the same population was used to help control confounding factors (Hartung & Touchette, 2009) and rule out any secular trends that may have resulted from being enrolled in the Basic BScN nursing program (DiCenso, Guyatt, & Ciliska, 2005).

Participants

McMaster University's collaborative baccalaureate nursing program consists of three sites: McMaster University, Mohawk College, and Conestoga College. McMaster University and Mohawk College nursing students study at the McMaster University campus in Hamilton and Conestoga College students study in Kitchener, Ontario. Hence, a study sample was drawn from only McMaster University and Mohawk College due to convenience of location. Baccalaureate nursing students enrolled in first and second year were recruited for the study for several reasons which included: (1) the focus of the study was on novice nursing students; (2) junior baccalaureate nursing students are new to problem-based learning and are not as immersed in this pedagogy as senior nursing students may be; and (3) the focus of learning in first and second year is on identification and interpretation while the focus of senior years is on application and evaluation (McMaster University BScN Handbook, 2012).

Sample Size

Based on results from the pilot study, the preliminary sample size was estimated using the following formula (Chow, Shao & Wang, 2008; Cleophas, Zwinderman, Cleophas & Cleophas, 2009; Gogtay, 2010):

$$n = \frac{2\left(Z_{\alpha/2} + Z_{\beta}\right)^2 \sigma^2}{d^2}$$

In this formula:

1. $Z_{\alpha/2}$ represents the critical value (e.g., 1.96) on a normal probability graph which the test statistic is unlikely to lie beyond. This means that the probability of falsely rejecting the true null hypothesis is 5% (i.e., α = 0.05).
2. Z_{β} represents the value (e.g., 0.842) on a normal probability graph which is used to specify the study power. This means that the probability of failing to reject the false null hypothesis is 20% (i.e., β = 0.2).
3. σ^2 represents the assumed pooled variance in the control and experimental groups.
4. d^2 represents the expected difference between control and experimental groups.

Since the pilot study did not have a control group, participants were used both as control group (i.e., pre-test data) and experimental group (post-test experimental data). Estimations of *d* were based on expert feedback. Sample size calculations for outcome variables (hypothesis accuracy, expertise and breadth) are presented in Appendix 3. Based on these calculations, hypothesis breadth generated the largest sample size requirement of 24 participants per group. Therefore, this figure was selected as a guideline for the study sample of 48 participants (control=24, treatment=24).

Recruitment

Through the internet, word of mouth, posters, and announcements outside of class time, all first and second year baccalaureate nursing students who attended classes at the McMaster University campus were invited to participate in the study. Only students who voluntarily came forward and met the eligibility criteria were recruited in the study. Key faculty members at the McMaster School of Nursing were also consulted from time to time in order to improve recruitment efforts and enrolment of participants in the study.

Enrolment and Assignment

Enrolment in the study was voluntary. Students interested in participating in the study were requested to contact the researcher. Students were assigned to either the control or intervention group based on the day in which they studied the problem-based learning course. For example, first year baccalaureate nursing students at the McMaster University site study the problem-based learning course on either Monday or Friday. Therefore, participants who attend the Monday class were assigned to the control group and those attending the Friday class were assigned to the treatment group.

The process of assignment was similar for second year baccalaureate nursing students and those students belonging to the Mohawk College site. In all cases, attempts were made to assign participants to either the control or treatment group based on the day they studied the problem-based learning course. However, in some circumstances, some participants were assigned to a group that was different from the day of their class due to reasons that were difficult to control. These included participants that: (a) were roommates of another participant, (b) were friends with another participant, or (c) were only available at a time during which the contrary study session was taking place.

Inclusion and Exclusion

Only nursing students enrolled in first and second year of the four-year Basic BScN program who were attending High School prior to commencing the BScN program were included in the study. Transfer students, students in the accelerated or RPN-to-BScN streams, and students with prior post-secondary education were excluded. The inclusion and exclusion criteria were made explicit to potential participants during recruitment.

Setting and Process

For students' convenience, study rooms were booked on campus through the university's room booking service. Flexible scheduling of study sessions was also used depending on students' availability. The study included a 105-minute pre-post study session and a 15-minute follow-up session one week later (Appendix 4).

Incentives

According to the Higher Education Quality Council of Ontario (Cyr, Childs & Elgie, 2013), offering incentives corresponds well with the social exchange theory of attracting participants. In their report, Cyr et al. indicate that various incentives have been used to recruit students (e.g., food, a small education-related item, a small course grade adjustment, monetary reimbursement, lottery draw for a larger prize, etc.); however, there is no standard practice for choosing which incentive to offer participants. The authors discuss Singer and Bossarte (2006) who highlight that money as an incentive is more effective than non-cash incentives. As a guide, Cyr et al. also mention the work of Dickert and Grady (1999) and Grady (2001) who emphasize that financial incentives should be similar to a working wage.

Students received $20 for their participation (i.e., $10 at the first study session and $10 at the follow-up session). The amount was split into two sessions in order to prevent drop-out of participants. Since a two-hour commitment was required from participants, a $10 per hour rate was selected based on current wage trends. However, there were challenges in student recruitment which were expected. To boost participation, several announcements were made that if students signed up on the day of the announcement, they will receive chocolate bars. This strategy was very effective, especially due to technology which allowed students to touch base with the researcher within seconds of the announcements.

Intervention

Hypothesis Education (20-minute)

Participants were provided with hypothesis education at the beginning to ensure they developed a common understanding of what a hypothesis is. They were then tested through a picture exercise in which they were required to generate hypotheses for various pictures that presented a different care situation.

Pre-Test (15-minute)

After education, participants were asked to complete the pre-test exercise (Appendix 5). The pre-test exercise consisted of a scenario and two questions – the first asking to generate issues in the care scenario and the second asking to generate hypotheses to explain the care scenario. After this, participants received the intervention depending on the study group to which they were assigned (i.e., control participants received discussion + standard training, and treatment participants received abductive reasoning + standard training).

Control Group: Discussion + Standard Training (40-minute)

Control participants participated in a 20-minute discussion on the process of hypothesis generation which was followed by a 20-minute group activity in which participants worked together on a given care scenario to generate hypotheses as a group (Appendix 6).

Treatment Group: Abductive Reasoning + Standard Training (40-minute)

Intervention participants participated in a 20-minute training on how to apply abductive reasoning. This was followed by a 20-minute group exercise in which participants worked together to apply abductive reasoning skills to a given care scenario in order to generate hypotheses as a group (Appendix 7). According to Kinchin, et al. (2005), visual models involving concepts and their relationships with one another (e.g., concept maps) can be taught in 20 minutes in a collaborative setting. Therefore, this guideline was used for the development of causal models based on abductive reasoning.

Immediate Post-Test (15-minute)

Immediately following training, all participants completed a 15-minute post-test exercise (Appendix 8). This was similar to the pre-test exercise and also consisted of a care scenario and two questions which requested participants to list issues and generate hypotheses to explain the care situation within the scenario. Participants in the treatment group were also given a blank piece of paper and encouraged to draw an abductive causal model to help their understanding of the care scenario. They were also given 15 minutes in order to ensure consistency between control and treatment groups. A 15-minute time frame was used since Kinchin et al. (2005) have found 15 minutes to be sufficient in allowing students to draw visual models of concepts and their relationships which they are learning.

One-Week Follow-Up Test (15-minute)

All participants were invited to complete a follow-up one week after the training. The follow-up consisted of another 15-minute test exercise (Appendix 9), which was similar to the pre-test and post-test exercises and also consisted of a care scenario with two questions requesting participants to list issues and generate hypotheses to explain the care scenario. A follow-up of one-week was used for several reasons. First, researchers exploring physicians' knowledge retention have indicated that physicians retain approximately half of their learning at 3 to 8 days with no significant retention at 55 days (Bell et al., 2008). Second, several experts of healthcare education have also used one-week (i.e., 7 days) as an appropriate follow-up for evaluating novice clinicians' abilities to diagnose and hypothesize about given care situations (Roediger & Karpicke, 2006; Woods, Brooks & Norman, 2007).

Finally, personal communication with Dr. Henry Roediger III (Professor, Department of Psychology & Memory Lab, Washington University in St. Louis) and Dr. Nicole Woods (Education Scientist, The Wilson Centre & Director, Education Evaluation, Department of Surgery, University of Toronto) indicated that one-week follow up was sufficient for health sciences university students for two main reasons, practicality and memory. According to Roediger, if students are able to come at a certain time on a particular day one week, they are more likely to be free the same time and day the following week. Moreover, a one-week delay facilitates recruitment, scheduling, and retaining student participants in a study (Woods). In terms of memory, both experts (Roediger; Woods) viewed one week as a long enough period to allow students to forget what they learned, but not long enough that they will forget to return for follow-up.

Content Validity

Content validity is the extent to which an instrument adequately covers a content area (Loiselle, Profetto-McGrath, Polit, & Beck, 2011). In this study, content validity of specific care scenarios was assessed to determine whether or not they were clear, sufficient to assess hypothesis generation abilities of novice nursing students, and appropriate for the target population (i.e., first and second year baccalaureate nursing students). In addition to this, picture-based issues and scenario-derived issues and hypotheses were also assessed for content validity. To obtain content validity of these instruments, five content experts and two content verifiers were utilized, all of whom were involved in teaching within the Basic four-year BScN program at the McMaster University School of Nursing.

Scenario Development

Three care scenarios were developed in a way which ensured that they were appropriate for the target population. To do this, all three care scenarios were created based on curriculum themes of first year (i.e., wellness and health promotion). This warranted that participants enrolled in first year would not feel overwhelmed with care scenarios of advanced complexity, and participants enrolled in second year would be familiar with first year themes. Health issues pertaining to teenage girls of junior High School age were discussed in the care scenarios. This approach ensured consistency of a similar theme (i.e., health and well being of adolescent females) and level of difficulty across all scenarios. The scenarios were also designed in such a way that they would allow participants to generate at least three issues and one hypothesis per biopsychosocial (biological, psychological, social) category.

Content Experts

Content validity was sought through the use of experts. According to Polit and Beck (2012), 5 to 8 experts are needed to provide content validity. Hence, 5 experts were selected based on their experience and involvement in either one or a combination of: (1) teaching in the undergraduate nursing education within the ‘Person-Based Learning within Problem-Based Learning Model’ at McMaster University School of Nursing; (2) the development of care scenarios, (3) instrument development, and (4) recommendation from course planners of first and second year. All five experts were faculty members at the McMaster School of Nursing and were currently teaching in the Basic Stream of the four-year baccalaureate nursing program. Of these, two were chairs (first year and second year) and one was a course planner while two were in leadership positions within the undergraduate nursing program.

Experts were sent a packet which included: (a) cover letter; (b) background information about the study and target population; (c) reviewer instructions; (d) measurement tools (picture scenarios with list of possible hypotheses, and written care scenarios with list of possible issues and hypotheses); and (e) questionnaires specific to each of the measurement tools. Experts were requested to use a 4.0 scale recommended by Polit and Beck (2012) in order to rate:

1. The ideal issue relationships in the picture exercise scenarios.
2. Care scenarios based on clarity, sufficiency and appropriateness for the target population.
3. The ideal list of issues for each care scenario based on breadth categories.
4. The ideal list of relationships for each care scenario based on breadth categories.
5. Consistency in similarity and breadth among scenarios.

Content Verifiers

Additional items were suggested by content experts for the various pictures and scenarios. Since the recommendations by the content experts varied, the newly generated items required further validation. The literature recommends the use of either new content experts or a mix of both previous and new experts (Polit & Beck, 2012). Hence, new experts were sought to verify additions to the content in order to avoid bias from previous experts. To do this, one faculty member who teaches in the undergraduate BScN program and one full-time graduate student enrolled in the PhD nursing program who also has teaching responsibilities were approached and they both agreed to contribute as content verifiers.

Content Validity Index

The content validity of an instrument is based on subjective judgment and is measured through a content validity index (CVI) which indicates the extent to which field experts agree about the items being measured (Loiselle et al., 2011). To ensure the validity of care scenarios, a CVI was used. The CVI was based on responses (issues and hypotheses) for picture-based and scenario-based exercises, and were categorized according to the various breadth categories (i.e., biological, psychological, social). As per Polit and Beck (2012), *content experts* were asked to rate the relevance of each issue per category per scenario from 1 (not relevant) to 4 (highly relevant). *Content verifiers*, on the other hand, were only asked to rate the relevance of items which required further verification.

In accordance with Polit and Beck's (2012) method, for each item, the *item CVI* (I-CVI) was computed as the number of experts rating 3 or 4, divided by the number of total experts who responded to the item. This provided the proportion of experts who were in agreement about relevance of a certain item. Since the ideal I-CVI is considered to be .78 or higher (Polit & Beck), only those items with an I-CVI of .78 or greater were included in the ideal list of responses. The I-CVI score for one scenario was rated below what was desired by one expert. With clarification and feedback, the order of the sentences in that particular scenario was altered to ensure better flow in reading. Upon this revision, the content expert revised her rating which resulted in an acceptable I-CVI score.

Data Collection

Test Questionnaires

Three test questionnaires were prepared to collect data (e.g., pre-test, post-test, follow-up test). For tracking purposes, participants were requested to write the last four digits of their university I.D. number which is unique to each student. All test questionnaires were numbered based on allocation – i.e., at the end of the session, a 1 for control group participants and a 2 for treatment group participants were indicated. After each study session, collected data were placed in an envelope which was dated, labelled with the experimental group (control or intervention) and placed in a designated place.

Demographic Data

Specific demographic data were collected. These consisted of:

Gender. While medical education research has found no significant gender differences in learning style preferences of novice medical students (Choudhary, Dullo & Tandon, 2011), there are, however, differences in the amount of information males and females can process. Females are known to have superior ability in multitasking which allows them to be more advantageous in having a stronger working memory than that of men. This allows them to process more information at a given time (Knox, Seth, McElveen, Bergstein & Longo, 2003).

Age. Since exposure to various life experiences increases with age, it has been found that individuals with more experience tend to possess superior abilities in retaining information when compared to those with little experience. While the number of retained items is similar among those with a lot of experience and those with very little experience, there is a correlation between the amount of experience an individual has and the amount of details he or she can retain and recall (Ericsson & Stasewski, 1989).

Program site. High School grades are known to predict academic performance in university (Winter & Dodou, 2011). Since High School entrance averages for the collaborative BScN program at McMaster University vary from site to site, data was collected on site. While applicants with a High School average between 88-91% are anticipated at the McMaster University site, a 75% High School is acceptable at the Mohawk College site (McMaster University, 2015; Mohawk College, 2015).

Year in program. Since baccalaureate nursing students from both first and second year were enrolled in the study, data was collected on their year in program. Since second year participants would have obtained a larger knowledge base after completing first year, they would ideally possess more prior knowledge to apply to the care scenarios presented in the study tests and, as a result, may perform differently than participants in first year.

Extracurricular training and service experience. Extracurricular activities (e.g., sports and non-sports) have been found to positively influence academic performance (Guest & Schneider, 2003; Marsh & Kleitman, 2002). Additionally, volunteering has also been found to have positive effects on academic performance of students (Hinck & Brandell, 1999). Moreover, Darling, Caldwell and Smith (2005) found that student who do not participate in any extracurricular activities have the poorest grades, attitude towards school, and academic aspirations.

Analysis

Outcome Variables

Study outcomes were based on the work of Ingram et al. (1998), and consisted of: total responses, correct responses, accuracy, expertise, and breadth. Of these, hypothesis accuracy, expertise, and breadth were the three main outcome variables. Ingram et al. describe the three main outcomes along with others as:

Number of responses. This refers to the total number of responses (issues or hypotheses) generated by the study participant for a given scenario.

Correct responses. Participants' responses which match the experts' list of responses (issues or hypotheses) for a given scenario are classified as correct.

Accuracy. This refers to the number of correct responses (issues or hypotheses) per care scenario as a percentage of the total number of responses generated by the participant for that scenario. While accuracy is described as the number of correct responses in medical literature (Elstein et al., 1978), Ingram et al. (1998) describe it as the percentage of the total number of correct responses out of the total number of responses generated by the participant.

Expertise. This refers to the number of correct responses (issues or hypotheses) per care scenario as a percentage of all ideal responses verified by experts for that scenario. Within the problem-based learning context, Margetson (1997) refers to expertise as the ability to make *sound judgments* about a care situation and identify its most important features. Similarly, Rideout (2001a) refers to expertise as the ability to make *appropriate judgments* about the main features of a care situation. Hence, the term expertise was used to refer to the percentage of the total number of correct responses out of the total number of ideal responses generated by field experts.

Breadth. This refers to the number of biopsychosocial categories which participants' responses (issues or hypotheses) covered per care scenario. Ingram et al. (1998) along with Rideout (2001a) refer to breadth in similar ways. Rideout uses the biopsychosocial framework (i.e., biological, psychological, and social aspects of health) while Ingram et al., in addition to these three categories, also use developmental, spiritual/cultural, and political/economic categories. However, to be congruent with the person-based learning within problem-based learning pedagogy used at the McMaster School of Nursing (McMaster University BScN Handbook, 2012), only the biopsychosocial categories were used to assess breadth. Limiting breadth to three categories also simplified the rating process for the two raters. This also ensured better agreement between raters, which would have been difficult to achieve if they had rated participant responses against six or seven breadth categories.

Raters

Two raters rated all study test questionnaires. To minimize selection bias of raters, four names of potential raters were placed in RANDOM.ORG which is a free, web-based, true randomizing service. The first two candidates on the output list were approached and provided information on the study and their possible responsibilities as raters. Both candidates agreed to participate in the study. Phone conversations were held with each of them to answer any questions they had about their role as raters. Both raters were registered nurses who had obtained a Masters of Nursing from a university in the Greater

Toronto Area within the last three years, and were not associated with McMaster University or Hamilton Health Sciences in any way.

A three-hour training session was organized for both raters at a convenient location. They were asked to sign a contract to ensure that all study information would be kept confidential and all study material would be returned to the researcher once they had completed rating the study data. During the training, all study material was discussed and raters were provided with sample test questionnaires from the pilot study to practice on. Once both raters gave similar ratings to the sample questionnaires, they were provided with printed copies of the test questionnaires which they were asked to rate independently. Both raters remained blinded to the allocation of participants and did not have any contact with one another during the rating process.

Inter-Rater Reliability

Inter-rater reliability refers to the degree of agreement between two or more examiners who independently rate an attribute being measured (Polit & Beck, 2012). The inter-rater reliability between two raters was assessed by the intraclass correlation coefficient (ICC). When determining which type of ICC to use, it is important to first clarify how the raters were identified. If the same raters rate every participant, then a two-way model of ICC must be used. However, if different raters make ratings on each participant, a one-way random ICC should be used (Landers, 2013). Since the two raters in this study made ratings on every participant, a two-way random ICC was used.

When considering a two-way ICC model, one must determine whether it is more appropriate to use a two-way random or a two-way mixed approach to calculating the

ICC. To figure this out, it is important to further clarify how the raters were identified. If a sample of raters is used, then a two-way random model should be considered. However, if a population of raters is used, then a two-way mixed model must be used (Landers, 2013). Since a sample of raters was used in the study, the two-way random model for calculating the ICC was considered. The ICC was calculated using SPSS software for statistical analysis.

Analysis of Covariance

While analysis of variance (ANOVA) is used to determine a significant difference between the means of two or more independent groups (e.g., control and treatment groups), analysis of covariance (ANCOVA) goes a step further by allowing the researcher to statistically control for some confounding variables known as the covariates. In studies lacking experimental control through randomization, ANCOVA ensures that study findings are not the result of pre-existing group differences by adjusting for initial differences between the control and experimental groups. This generates results that reflect a more precise estimate of the effect of the intervention (Polit & Beck, 2012).

Since hypothesis generation abilities were measured before (pre-test), immediately after (post-test), and one-week after (follow-up) the intervention, pre-test scores at baseline were statistically controlled through ANCOVA. To do this, immediate post-test and one-week follow-up scores were set as dependent variables, control or treatment group status was set as the independent variable, and pre-test scores were set as covariates. The ANCOVA statistical test was conducted using SPSS and the test was repeated for all outcome variables for both issues and hypotheses.

Other Statistical Tests

In addition to ICC and ANCOVA, Chi-square test was used to compare demographic data between control and treatment groups. Independent *t*-test was also used to compare baseline (pre-test) data between control and treatment groups. While ANCOVA addressed the primary study objective and main study hypotheses (i.e., to compare hypothesis accuracy, expertise and breadth between control and treatment groups at different time points), a two-way ANCOVA was also used to test the secondary objective of the study (i.e., to compare hypothesis accuracy, expertise and breadth between first and second year participants). The difference in the two-way ANCOVA is that in addition to having one factor (i.e., one-way) as an independent variable, two factors (i.e., two-way) are used as independent variables. Therefore, in the two-way ANCOVA, study year (first or second) was also used as an independent variable in addition to group allocation (control or treatment).

Management

Data collection sheets were number-coded. Data were entered by a research assistant into SPSS in a computer with a password. During the data entry process, missing data was identified and raters were approached to fill incomplete ratings or to clarify ratings which had visible errors (e.g., 5 total hypotheses, of which 6 were correct). Once data entry was complete, the researcher verified 100% of the data to ensure no mistakes were made during the data entry process. This was done by going through each data value to ensure it was correctly entered. Once electronic data were verified, hard-copies of data were stored in a locked cupboard. The electronic data was then transferred into SPSS. In all of data entry steps, no identifying information was entered into the electronic files which ensured anonymity of participants. A statistician was consulted regularly for guidance in matters which involved statistical analyses and interpretation of statistical tests.

Ethics

To prepare for the role, the researcher completed an ethics certificate, attended an ethics workshop, and obtained ethical approval for the study from the Hamilton Integrated Research Ethics Board. Permission to conduct the research was also obtained from the Undergraduate Nursing Education Committee at McMaster University. A copy of the ethics approval is included in Appendix 10, and a copy of the consent form is included in Appendix 11. During the consent process, the study was explained to students. They were then given time to read the consent form and ask any questions. Once they read the consent form and agreed to participate in the study, they signed two copies of the consent form (one for them to keep and one for the researcher). Each consent form was also signed by a witness (i.e., another participant in the group) and the researcher.

To ensure participants did not have any questions, they were asked by the researcher if they had any questions about the study or the consent form before they signed the consent form. Any questions by the participants were answered. Participants were also told that participation in the study was voluntary and that they could withdraw from the study at any time without having to provide a reason why they wanted to withdraw. Moreover, participants were ensured that participation in the study would not affect their role as a student at the university. They were also ensured that no faculty member would be involved in rating the data collected. To avoid overlap with important academic events, all study-related announcements were made outside of class time and participants were not recruited during exam periods.

CHAPTER 4: RESULTS

A total number of 64 baccalaureate nursing students from the four-year Basic BScN program at McMaster University who met eligibility criteria participated in the study. Of these, 29 were in the control group while 35 were in the experimental group. Before completing the baseline test questionnaire, all 64 study participants received education on what a hypothesis is. This was followed by two different interventions. Control participants participated in a group discussion on the process of hypothesis generation, and applied their newly acquired learning to the presenting care scenario. Experimental participants, on the other hand, received abductive reasoning training and applied their newly acquired learning to the same presenting scenario as the control group. After the different interventions, both control and experimental participants completed an immediate post-test questionnaire. All participants were then asked to return in a week's time to complete the follow-up test questionnaire based on a new scenario.

To conduct the study, participants were divided into 15 groups of varying sizes depending on participants' time preference. A total of 30 study sessions were held (15 sessions for baseline and immediate post-test and 15 sessions for follow-up). In addition to these sessions, several study sessions had to be cancelled due to no show or prior notification from attendees. All group sessions took place in the McMaster University Medical Centre. Group size ranged from 2 to 13 participants (mean 4.3, median 3, mode 3, minimum 2, maximum 13). On one occasion, a small-group session for 6 participants was expanded to a large-group session to accommodate an additional 4 participants who had been unable to attend previously scheduled study sessions. Participants from this large group session also brought 3 additional participants, which further expanded the group size to 13 participants. This accounts for the maximum value of 13 in group size.

Follow-up was 100% with no missing data. There was one participant who completed follow-up at 21 days due to circumstances which were beyond the control of the researcher. Despite this, follow-up time generally varied from 7 to 8 days for the rest of the 63 participants (mean 7.3, median 7, mode 7, minimum 7, maximum 21).

Comparison between Control and Experimental Groups at Baseline

To ensure both control and experimental groups were similar, demographic variables between control and experimental participants were compared. These variables were age, gender, site, year in program, and extracurricular experience (Table 4.1).

Table 4.1: Frequency [n(%)] Demographics and Baseline Characteristics in Control and Experimental Groups

Variable	Control (N=29)	Experimental (N=35)	Overall (N=64)	P-value*
Age Group				0.811
17-18	11 (37.9)	15 (42.9)	26 (40.6)	
19-20	17 (58.6)	18 (51.4)	35 (54.7)	
21 or above	1 (3.4)	2 (5.7)	3 (4.7)	
Gender				0.492
Male	3 (10.3)	2 (5.7)	5 (7.8)	
Female	26 (89.7)	33 (94.3)	59 (92.2)	
Site				0.713
McMaster	25 (86.2)	29 (82.9)	54 (84.4)	
Mohawk	4 (13.8)	6 (17.1)	10 (15.6)	
Year in Program				0.911
First Year	17 (58.6)	21 (60.0)	38 (59.4)	
Second Year	12 (41.4)	14 (40.0)	26 (40.6)	
Extracurricular				0.892
Have experience	28 (96.6)	34 (97.1)	62 (96.9)	
Do not have experience	1 (3.4)	1 (2.9)	2 (3.1)	

* P-value based on Chi-square test

Demographic Variables

Age. Age was divided into three groups (i.e., 17-18, 19-20, and 21 and above). While 26 (40.6%) participants were in the 17-18 age group, 35 (54.7%) participants were in the 19-20 age group and 3 (4.7%) were in the third age group. In the 17-18 age group, 11 participants were in the control group and 15 in the experimental group. Respectively, these figures were 17 and 18 participants in the 19-20 age group, and 1 and 2 participants in the 21 and above age group. The Chi-square test revealed that there was no statistically significant difference between control and experimental participants in terms of age (p=0.811).

Gender. Both male and female baccalaureate nursing students participated in the study. Overall, there were 5 (7.8%) male and 59 (92.2%) female participants in the study. While there were 3 male participants in the control group, the experimental group consisted of 2 male participants. These figures for female participants were 26 in the control group and 33 in the experimental group. There was no statistically significant difference between the control and experimental groups in terms of gender (p=0.492).

Site. Baccalaureate nursing students from both the McMaster University site and the Mohawk College site participated in the study. However, 54 (84.4%) participants were from the McMaster University site while 10 (15.6%) were from the Mohawk College site. Of these participants, 25 McMaster and 4 Mohawk site participants were in the control group while 29 McMaster and 6 Mohawk site participants were in the experimental group. In terms of site, no statistically significant differences were found between control and experimental participants (p=0.713).

Year in program. Thirty eight (59.4%) first year and 26 (40.6%) second year baccalaureate nursing students participated in the study. While there were 17 first year participants in the control group, there were 21 first year participants in the experimental group. These figures for second year students were 12 in the control group and 14 in the experimental group. No statistically significant difference between the control and experimental groups was detected in terms of year of study (p=0.911).

Extracurricular experience. Most study participants had had some kind of extracurricular experience prior to participating in the study. While 62 (96.9%) participants had participated in extracurricular activities in the past, 2 (3.1%) participants had no extracurricular experience. Of the 2 participants that did not have extracurricular experience, one was in the control group and the other was in the experimental group. Of the remaining 62 participants with extracurricular experience, 28 were in the control group and 34 were in the experimental group. Using the Chi-square test, no statistically significant difference was detected between the control and experimental groups in terms of extracurricular experience (p=0.892).

Calculation of Inter-Rater Reliability

All study test questionnaires were rated by two raters. To ensure raters rated test questionnaires similarly (i.e., inter-rater reliability), the intraclass correlation coefficient (ICC) was calculated for each variable at baseline pre-test, immediate post-test and at one-week follow-up. The ICC value for each variable is presented in Table 4.2 and varies from 0.625 to 1.00 for all variables. However, for the outcome variables (i.e., hypothesis accuracy, hypothesis expertise, and hypothesis breadth), the ICC varied from 0.732 to

0.811 at baseline and 0.774 to 0.861 at one-week follow-up. As outlined by Polit and Beck (2012), most of these ICC values are more than the adequate and acceptable value of 0.70 for reliability coefficients. For analysis and reporting purposes, the mean of both raters was used.

Table 4.2: Inter-Rater Reliability Between Two Raters

	ICC			
Hypothesis Education	0.748			
		Baseline Pre-Test	*Immediate Post-Test*	*One-Week Follow-Up*
Issues				
Total		0.999	0.992	0.995
Correct		0.752	0.870	0.897
Accuracy		0.726	0.625	0.789
Expertise		0.752	0.870	0.897
Breadth		0.648	0.791	0.672
Hypotheses				
Total		1.00	0.999	0.996
Correct		0.732	0.745	0.861
Accuracy		0.741	0.835	0.774
Expertise		0.732	0.745	0.861
Breadth		0.811	0.671	0.801

Post-Hoc Power Analysis

The initial sample size for the study was based on a pre-post pilot study which was conducted in one session and did not entirely resemble the larger study which followed. For this reason, a post-hoc power analysis was calculated after the large study was conducted. A post-hoc power analysis is the retrospective power calculation of an observed value of the effect size (Lenth, 2007). The power of the study, which is calculated by the post-hoc power analysis, is known as the probability of correctly rejecting the null hypothesis in the study sample (Hulley, Cummings, Browner, Grady & Newman, 2007).

The Power and Sample Size Calculations software provided electronically by the Vanderbilt University Department of Biostatistics (2014) was used to carry out post-hoc power analyses for each of the three outcome variables (hypothesis accuracy, expertise and breadth). Based on the study results, the analysis was performed using the number of participants in the experimental group (i.e., 35), α of 0.05, the ratio between control and experimental participants (29:35=0.8286), the observed difference between the control and experimental groups, and the pooled standard deviation of the control and experimental group for each variable using the following formula (Daniel & Cross, 2013):

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

The power of the study was calculated to be: (a) 0.503 for hypothesis accuracy, (b) 0.808 for hypothesis expertise, and (c) 0.862 for hypothesis breadth. This showed that the population means of the control and experimental groups were different in hypothesis accuracy with probability (power) of 0.503, while the groups were different in hypothesis expertise and hypothesis breadth with probabilities (powers) of 0.808 and 0.862 respectively. These calculations are presented in Table 4.3.

Table 4.3: Power Calculations based on Study Results

	Accuracy	Expertise	Breadth
α	0.05	0.05	0.05
δ (Observed Mean Difference)	0.065	0.052	0.613
σ (Pooled Standard Deviation)	0.129	0.072	0.788
m **($n_{control}$ / $n_{experimental}$)**	0.829	0.829	0.829
Power	0.503	0.808	0.862

Education on Hypothesis

At the start of the first study session, all participants were provided with education on what a hypothesis is. Based on Miller's (1990) pyramid model of competence, the purpose of this education was to assess participants' knowledge of hypothesis, provide education to create a similar understanding of hypothesis among all participants, and to assess whether participants know how to generate hypotheses through the use of picture scenarios. Based on the mean score of two raters who rated participants' hypotheses for picture scenarios, the mean hypothesis knowledge scores were 1.83 (SD 0.72) out of 3.0 for control participants and 1.90 (SD 0.86) out of 3.0 for experimental participants. The minimum and maximum scores for both groups were 0.0 and 3.0. No statistically significant difference was detected between the control and experimental groups ($p=0.719$).

Outcome Variables

While there were three main outcome variables (i.e., hypothesis accuracy, hypothesis expertise and hypothesis breadth), ten outcome variables were measured. These outcome variables are reported as a mean value based on ratings of two raters in Table 4.4. Participants were asked to generate issues and hypotheses based on care scenarios. Issues (descriptive hypotheses) consisted of one-variable, while hypotheses (relational hypotheses) consisted of two variables (i.e., independent and dependent) which were linked together to show a relationship. First, the total number of issues and hypotheses generated by the participant were measured. Second, the total number of issues and hypotheses that matched the expert list of responses (i.e., correct issues, correct hypotheses) were measured.

Following the calculation of the number of total and correct issues and hypotheses, the three study variables were calculated. The accuracy score was calculated by taking the number of correct responses and dividing by the number of total responses for issues and hypotheses (i.e., issue accuracy, hypothesis accuracy). Similarly, expertise was calculated by dividing the correct responses by the possible number of correct responses based on the expert list (i.e., issue expertise, hypothesis expertise). Finally, breadth was calculated based on the number of biopsychosocial categories (i.e., biological, psychological, social) which participants addressed in their responses. For example, if a participant's responses addressed 2 of the 3 categories, the participant received a score of 2 out of 3 in either issue breadth or hypothesis breadth.

Table 4.4: Comparison between Control and Experimental Groups at Different Time Points‡

	Control (N=29)				Experimental (N=35)				
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.	P-value
Hypothesis Education*									
	1.83	0.723	0	3.00	1.90	0.856	0	3.00	0.719
Baseline Pre-Test*									
Issues									
Total	6.52	2.09	3.00	11.00	6.34	2.52	3.00	15.00	0.767
Correct	3.90	1.55	1.50	8.00	3.83	1.13	2.00	6.00	0.840
Accuracy †	0.62	0.20	0.21	1.00	0.65	0.18	0.25	1.00	0.531
Expertise †	0.32	0.13	0.13	0.67	0.32	0.09	0.17	0.50	0.840
Breadth †	2.17	0.56	1.00	3.00	2.31	0.49	1.50	3.00	0.280
Hypotheses									
Total	3.14	1.73	1.00	8.00	2.34	1.35	1.00	5.00	**0.043**
Correct	0.29	0.54	0	2.00	0.24	0.37	0	1.00	0.663
Accuracy	0.11	0.19	0	0.50	0.11	0.21	0	1.00	0.983
Expertise	0.03	0.05	0	0.20	0.02	0.04	0	0.10	0.663
Breadth	0.47	0.87	0	2.50	0.47	0.74	0	2.00	0.977
Immediate Post-Test**									
Issues									
Total	7.45	2.08	3.00	12.00	11.79	2.80	7.00	19.50	**<0.001**
Correct	4.97	1.72	1.50	8.00	7.59	1.71	5.00	12.00	**<0.001**
Accuracy	0.67	0.16	0.22	0.92	0.65	0.09	0.50	0.85	0.531
Expertise	0.38	0.13	0.12	0.62	0.58	0.13	0.38	0.92	**<0.001**
Breadth	2.40	0.60	1.00	3.00	2.67	0.47	1.50	3.00	0.061
Hypotheses									
Total	4.07	1.93	1.00	9.00	6.21	2.87	1.00	13.00	**<0.001**
Correct	0.69	0.70	0	2.00	1.04	0.83	0	3.00	0.079
Accuracy	0.14	0.15	0	0.40	0.19	0.19	0	1.00	0.290
Expertise	0.07	0.07	0	0.20	0.10	0.08	0	0.30	0.079
Breadth	0.66	0.67	0	2.50	1.04	0.75	0	3.00	**0.036**
One-Week Follow-up**									
Issues									
Total	7.55	2.11	4.00	11.00	11.41	2.50	7.00	19.00	**<0.001**
Correct	4.88	1.53	2.00	7.50	7.09	1.12	5.00	10.00	**<0.001**
Accuracy	0.66	0.17	0.30	1.00	0.63	0.10	0.46	0.94	0.359
Expertise	0.41	0.13	0.17	0.63	0.59	0.09	0.42	0.83	**<0.001**
Breadth	2.74	0.46	1.50	3.50	2.86	0.33	1.50	3.00	0.310
Hypotheses									
Total	3.84	1.51	2.0	8.0	6.11	4.13	1.00	25.00	**0.001**
Correct	0.21	0.39	0	1.50	0.67	0.82	0	3.00	**0.006**
Accuracy	0.06	0.10	0	0.33	0.12	0.15	0	0.50	**0.050**
Expertise	0.02	0.04	0	0.17	0.07	0.09	0	0.33	**0.006**
Breadth	0.26	0.47	0	1.50	0.87	0.97	0	3.00	**0.003**

* P-value based by Independent T-test, ** P-value based on ANCOVA adjusted for baseline pre-test
† Accuracy and expertise were out of 1.0 while breadth was out of 3.0, ‡ Scores based on mean of two raters

Total Issues

Control participants generated a mean of 6.52 (SD 2.09) total issues at baseline while experimental participants generated 6.34 (SD 2.52) total issues. The minimum and maximum number of total issues generated by the control group was 3 and 11, while the minimum and maximum number for the experimental group was 3 and 15. There was no statistically significant difference in total number of issues generated between the control and experimental groups at baseline ($p=0.767$). This indicated that control and experimental groups generated a similar number of total issues at baseline.

At immediate post-test, control participants generated a mean of 7.45 (SD 2.08) total issues while experimental participants generated 11.79 (SD 2.80) total issues. The minimum and maximum number of issues generated by the control group was 3 and 12, while the minimum and maximum number for the experimental group was 7 and 19.5. A statistically significant difference was noted between the control and experimental participants based on the number of total issues generated ($p<0.001$). This indicated that the experimental participants generated significantly more issues than control participants.

While control participants generated a mean of 7.55 (SD 2.11) total issues at one-week follow-up, experimental participants generated a mean of 11.41 (SD 2.50) total issues. The minimum and maximum number of issues generated was 4 and 11 for the control group and 7 and 19 for the experimental group. Similar to the immediate post-test, a statistically significant difference in the number of total issues was noted between the control and experimental participants at one-week follow-up ($p<0.001$). This showed that the experimental group generated significantly more issues than the control group.

Correct Issues

Control participants generated a mean of 3.90 (SD 1.55) correct issues at baseline while experimental participants generated 3.83 (SD 1.13) correct issues. The minimum and maximum number of correct issues generated was 1.5 and 8 for the control group and 2 and 6 for the experimental group. There was no statistically significant difference in the number of correct issues generated by the control and experimental groups at baseline ($p=0.840$). This indicated that control and experimental groups generated a similar number of correct issues at baseline.

At immediate post-test, control participants generated a mean of 4.97 (SD 1.72) correct issues while experimental participants generated 7.59 (SD 1.71) correct issues. The minimum and maximum number of correct issues generated by the control group at immediate post-test was 1.5 and 8, while the experimental group generated 5 and 12 respectively. In terms of correct issues, a statistically significant difference was noted between the control and experimental groups ($p<0.001$). This indicated that the experimental participants generated significantly more correct issues than control participants.

While control participants generated a mean of 4.88 (SD 1.53) correct issues at one-week follow-up, experimental participants generated a mean of 7.09 (SD 1.12) correct issues. The minimum and maximum number of issues generated was 2 and 7.5 for the control group and 5 and 10 for the experimental group. Similar to the immediate post-test, a statistically significant difference in the number of correct issues was also noted between the control and experimental participants at one-week follow-up ($p<0.001$). This showed that the experimental group generated significantly more correct issues than the control group.

Issue Accuracy

Mean score for issue accuracy among control participants was 0.62 (SD 0.20) at baseline while it was 0.65 (SD 0.18) among experimental participants. The minimum and maximum issue accuracy score for control participants was 0.21 and 1.00, while the minimum and maximum issue accuracy score for experimental participants was 0.25 and 1.00. There was no statistically significant difference in issue accuracy score between the control and experimental groups at baseline ($p=0.531$). This indicated that control and experimental groups obtained similar issue accuracy scores at baseline.

At immediate post-test, the mean issue accuracy score among control participants was 0.67 (SD 0.16) and among experimental participants was 0.65 (SD 0.09). The minimum and maximum issue accuracy scores were respectively 0.22 and 0.92 for control participants and 0.50 and 0.85 for experimental participants. No statistically significant difference was noted between the control and experimental participants based on issue accuracy ($p=0.531$). This indicated that control and experimental groups obtained similar issue accuracy scores at immediate post-test.

The mean issue accuracy score at one-week follow-up was 0.66 (SD 0.17) for control participants and 0.63 (SD 0.10) for experimental participants. The minimum and maximum issue accuracy scores were 0.30 and 1.00 for control participants and 0.46 and

0.94 for experimental participants. Similar to the immediate post-test, no statistically significant difference in the issue accuracy score was noted between the control and experimental participants at one-week follow-up (p=0.359). This showed that both control and experimental participants scored similarly on issue accuracy at one-week follow-up.

Issue expertise

Mean issue expertise score at baseline was 0.32 (SD 0.13) for control participants and 0.32 (SD 0.09) for experimental participants. The minimum and maximum issue expertise scores were respectively 0.13 and 0.67 for control participants and 0.17 and 0.50 for experimental participants. There was no statistically significant difference in issue expertise score between the control and experimental groups at baseline (p=0.840). This indicated that control and experimental groups obtained similar issue expertise scores at baseline.

At immediate post-test, the mean issue expertise score among control participants was 0.38 (SD 0.13) and among experimental participants was 0.58 (SD 0.13). The minimum and maximum issue expertise scores were respectively 0.12 and 0.62 for control participants and 0.38 and 0.92 for experimental participants. A statistically significant difference was noted between the control and experimental participants based on issue expertise (p<0.001). This indicated that the experimental group scored significantly better than the control group on issue expertise at immediate post-test.

The mean expertise score at one-week follow-up was 0.41 (SD 0.13) for control participants and 0.59 (SD 0.09) for experimental participants. The minimum and maximum issue expertise scores were 0.17 and 0.63 for control participants and 0.42 and

0.83 for experimental participants. Similar to the immediate post-test, a statistically significant difference in the issue expertise score was noted between the control and experimental participants at one-week follow-up ($p<0.001$). This showed that the experimental group scored significantly better than the control group on issue expertise at one-week follow-up.

Issue Breadth

The mean group score for issue breadth was 2.17 (SD 0.56) at baseline for control participants and 2.31 (SD 0.49) for experimental participants. The minimum and maximum issue breadth scores were respectively 1.00 and 3.00 for control participants and 1.50 and 3.00 for experimental participants. There was no statistically significant difference in issue breadth score between the control and experimental groups at baseline ($p=0.280$). This indicated that control and experimental groups obtained similar issue breadth scores at baseline.

At immediate post-test, the mean issue breadth score among control participants was 2.40 (SD 0.60) and among experimental participants was 2.67 (SD 0.47). The minimum and maximum issue breadth scores were respectively 1.00 and 3.00 for control participants and 1.50 and 3.00 for experimental participants. No statistically significant difference was noted between the control and experimental participants based on issue breadth ($p=0.061$). This indicated that the control and experimental groups scored similarly on issue breadth at immediate post-test.

The mean breadth score at one-week follow-up was 2.74 (SD 0.46) for control participants and 2.86 (SD 0.33) for experimental participants. The minimum and maximum issue breadth scores were 1.50 and 3.50 for control participants and 1.50 and 3.00 for experimental participants. Similar to baseline and immediate post-test, no statistically significant difference in the issue breadth score was noted between the control and experimental participants at one-week follow-up (p=0.310). This showed that the control and experimental groups scored similarly on issue breadth at one-week follow-up.

Total Hypotheses

Control participants generated a mean of 3.14 (SD 1.73) total hypotheses at baseline while experimental participants generated 2.34 (SD 1.35) total hypotheses. The minimum and maximum number of total hypotheses generated by the control group was 1 and 8, while the minimum and maximum number for the experimental group was 1 and 5. There was statistically significant difference in total number of hypotheses generated between the control and experimental groups at baseline (p=0.043). This indicated that control participants generated significantly more hypotheses than the experimental participants at baseline.

At immediate post-test, control participants generated a mean of 4.07 (SD 1.93) total hypotheses while experimental participants generated 6.21 (SD 2.87) total hypotheses. The minimum and maximum number of hypotheses generated by the control group was 1 and 9, while the minimum and maximum number for the experimental group was 1 and 13. A statistically significant difference was noted between the control and experimental participants based on the number of total hypotheses generated (p<0.001). This indicated that the experimental participants generated significantly more hypotheses than control participants.

While control participants generated a mean of 3.84 (SD 1.51) total hypotheses at one-week follow-up, experimental participants generated a mean of 6.11 (SD 4.13) total hypotheses. The minimum and maximum number of hypotheses generated was 2 and 8 for the control group and 1 and 25 for the experimental group. Similar to the immediate post-test, a statistically significant difference in the number of total hypotheses was noted between the control and experimental participants at one-week follow-up (p=0.001). This showed that the experimental group generated significantly more hypotheses than the control group.

Correct Hypotheses

Control participants generated a mean of 0.29 (SD 0.54) correct hypotheses at baseline while experimental participants generated 0.24 (SD 0.37) correct hypotheses. The minimum and maximum number of correct hypotheses generated was 0 and 2 for the control group and 0 and 1 for the experimental group. There was no statistically significant difference in the number of correct hypotheses generated by the control and experimental groups at baseline (p=0.663). This indicated that control and experimental groups generated a similar number of correct hypotheses at baseline.

At immediate post-test, control participants generated a mean of 0.69 (SD 0.70) correct hypotheses while experimental participants generated 1.04 (SD 0.83) correct hypotheses. The minimum and maximum number of correct hypotheses generated by the control group at immediate post-test was 0 and 2, while the experimental group generated 0 and 3 respectively. In terms of correct hypotheses, no statistically significant difference was noted between the control and experimental groups (p=0.079). This indicated that the control and experimental participants generated a similar number of correct hypotheses at immediate post-test.

While control participants generated a mean of 0.21 (SD 0.39) correct hypotheses at one-week follow-up, experimental participants generated a mean of 0.67 (SD 0.82) correct hypotheses. The minimum and maximum number of hypotheses generated was 0 and 1.5 for the control group and 0 and 3 for the experimental group. While both control and experimental participants had generated a similar number of correct hypotheses at baseline and at immediate post-test, a statistically significant difference in the number of correct hypotheses was noted between the two groups at one-week follow-up (p=0.006). This showed that the experimental group generated significantly more correct hypotheses than the control group.

Hypothesis Accuracy

Mean score for hypothesis accuracy among control participants was 0.11 (SD 0.19) at baseline while it was 0.11 (SD 0.21) among experimental participants. The minimum and maximum hypothesis accuracy score for control participants was 0 and 0.50, while the minimum and maximum hypothesis accuracy score for experimental participants was 0 and 1.00. There was no statistically significant difference in hypothesis accuracy score between the control and experimental groups at baseline (p=0.983). This indicated that control and experimental groups obtained similar hypothesis accuracy scores at baseline.

At immediate post-test, the mean hypothesis accuracy score among control participants was 0.14 (SD 0.15) and among experimental participants was 0.19 (SD 0.19).

The minimum and maximum hypothesis accuracy scores were respectively 0 and 0.40 for control participants and 0 and 1.00 for experimental participants. No statistically significant difference was noted between the control and experimental participants based on hypothesis accuracy (p=0.290). This indicated that control and experimental groups obtained similar hypothesis accuracy scores at immediate post-test.

The mean hypothesis accuracy score at one-week follow-up was 0.06 (SD 0.10) for control participants and 0.12 (SD 0.15) for experimental participants. The minimum and maximum hypothesis accuracy scores were 0 and 0.33 for control participants and 0 and 0.50 for experimental participants. While both control and experimental participants had similar hypothesis accuracy scores at baseline and at immediate post-test, a statistically significant difference in the hypothesis accuracy score was noted between the two groups at one-week follow-up (p=0.050). This showed that the experimental group had significantly higher hypothesis accuracy than the control group at one-week follow-up.

Hypothesis Expertise

Mean hypothesis expertise score at baseline was 0.03 (SD 0.05) for control participants and 0.02 (SD 0.04) for experimental participants. The minimum and maximum hypothesis expertise scores were respectively 0 and 0.20 for control participants and 0 and 0.10 for experimental participants. There was no statistically significant difference in hypothesis expertise score between the control and experimental groups at baseline (p=0.663). This indicated that control and experimental groups obtained similar hypothesis expertise scores at baseline.

At immediate post-test, the mean hypothesis expertise score among control participants was 0.07 (SD 0.07) and among experimental participants was 0.10 (SD 0.08). The minimum and maximum hypothesis expertise scores were respectively 0 and 0.20 for control participants and 0 and 0.30 for experimental participants. In terms of hypothesis expertise, no statistically significant difference was noted between the control and experimental groups (p=0.079). This indicated that the control and experimental participants obtained similar hypothesis expertise score at immediate post-test.

The mean expertise score at one-week follow-up was 0.02 (SD 0.04) for control participants and 0.07 (SD 0.09) for experimental participants. The minimum and maximum hypothesis expertise scores were 0 and 0.17 for control participants and 0 and 0.33 for experimental participants. While both control and experimental participants had similar hypothesis expertise scores at baseline and at immediate post-test, a statistically significant difference in hypothesis expertise was detected between the two groups at one-week follow-up (p=0.006). This showed that the experimental group had significantly higher hypothesis expertise than the control group.

Hypothesis Breadth

Mean hypothesis breadth score at baseline was 0.47 (SD 0.87) for control participants and 0.47 (SD 0.74) for experimental participants. The minimum and maximum hypothesis breadth scores were respectively 0 and 2.50 for control participants and 0 and 2.00 for experimental participants. There was no statistically significant difference in hypothesis breadth score between the control and experimental groups at baseline (p=0.977). This indicated that control and experimental groups obtained similar hypothesis breadth scores at baseline.

At immediate post-test, the mean hypothesis breadth score among control participants was 0.66 (SD 0.67) and among experimental participants was 1.04 (SD 0.75). The minimum and maximum hypothesis breadth scores were respectively 0 and 2.50 for control participants and 0 and 3.00 for experimental participants. A statistically significant difference was noted between the control and experimental participants based on hypothesis breadth (p=0.036). This indicated that experimental participants scored significantly higher on hypothesis breadth compared to control participants at immediate post-test.

The mean breadth score at one-week follow-up was 0.26 (SD 0.47) for control participants and 0.87 (SD 0.97) for experimental participants. The minimum and maximum hypothesis breadth scores were 0 and 1.50 for control participants and 0 and 3.00 for experimental participants. Similar to immediate post-test, a statistically significant difference in the hypothesis breadth score was noted between the control and experimental participants at one-week follow-up (p=0.003). This showed that the experimental group scored significantly higher on hypothesis breadth at one-week follow-up than the control group.

Summative Assessment of Study Objectives

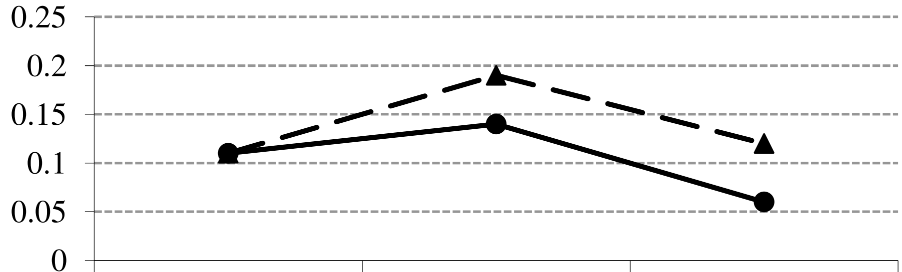
The main study objective was to examine the effect of abductive reasoning training on hypothesis generation abilities (accuracy, expertise, breadth) of first and second year baccalaureate nursing students. Study results indicated that compared to the control group, the experimental group, which was given the abductive reasoning training, had significantly higher scores on hypothesis breadth immediately after receiving the training (p<0.036). At one-week follow-up, experimental participants had significantly higher scores on all three study outcomes when compared to the control group. These abilities included hypothesis accuracy (p=0.050), hypothesis expertise (p=0.006), and hypothesis breadth (p=0.003). In addition to these, the experimental group also had significantly higher number of total hypotheses (p=0.001) and correct hypotheses (p=0.006) when compared to the control group at one-week follow-up (Table 4.5).

Table 4.5: Comparison of Participants based on Group and Year in Program

	Group					Year				
	Control (n=29)		Experimental (n=35)		P-value*†	First Year (n=38)		Second Year (n=26)		P-value*‡
	Mean	SD	Mean	SD		Mean	SD	Mean	SD	
Accuracy	0.06	0.10	0.12	0.15	**0.058**	0.10	0.14	0.09	0.13	**0.714**
Expertise	0.02	0.04	0.07	0.09	**0.007**	0.05	0.08	0.05	0.07	**0.873**
Breadth	0.26	0.47	0.87	0.97	**0.002**	0.53	0.74	0.69	0.97	**0.590**

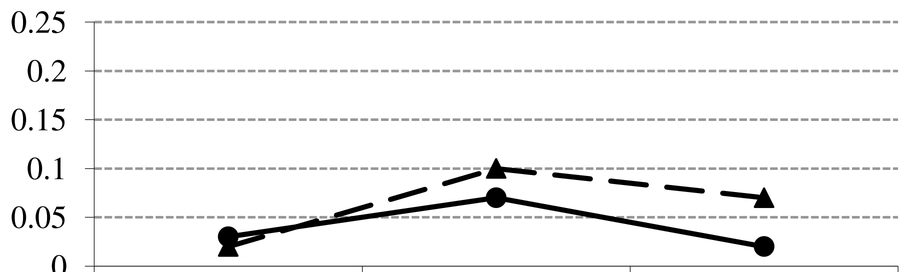
* P-value based on Two-Way ANCOVA
† Group p-value indicates difference between control and experimental groups
‡ Year p-value indicates difference between first and second year participants

Graphs illustrating observed trends between control and experimental participants are presented in Figure 4.1 for hypothesis accuracy, Figure 4.2 for hypothesis expertise, and Figure 4.3 for hypothesis breadth. These figures show that while both control and experimental participants performed similarly at baseline, there were significant differences in their performance at one-week follow-up.



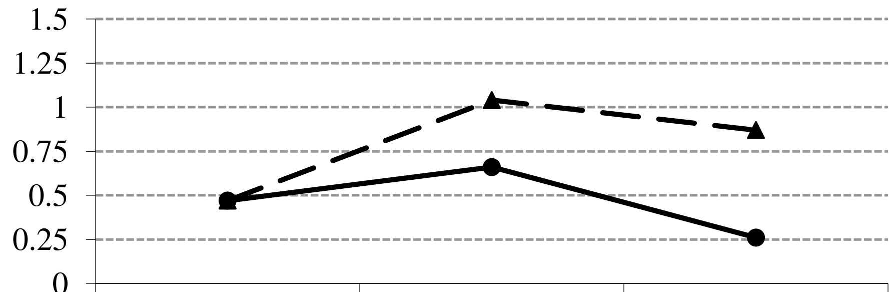
	Baseline Pre-Test (Mean)	Immediate Post-Test (Mean)	One-Week Follow-Up (Mean)
Control	0.11	0.14	0.06
Experimental	0.11	0.19	0.12

Figure 4.1. Comparison of hypothesis accuracy between control and experimental groups at different time points.



	Baseline Pre-Test (Mean)	Immediate Post-Test (Mean)	One-Week Follow-Up (Mean)
Control	0.03	0.07	0.02
Experimental	0.02	0.1	0.07

Figure 4.2. Comparison of hypothesis expertise between control and experimental groups at different time points.



	Baseline Pre-Test (Mean)	Immediate Post-Test (Mean)	One-Week Follow-Up (Mean)
Control	0.47	0.66	0.26
Experimental	0.47	1.04	0.87

Figure 4.3. Comparison of hypothesis breadth between control and experimental groups at different time points.

The second study objective was to compare study outcomes between first year and second year control and experimental participants. The purpose of this comparison was to determine whether or not the improvement between control and experimental groups was, in fact, due to the participation of second year students. However, no statistically significant differences were detected between first year and second year participants in hypothesis accuracy (p=0.714), hypothesis expertise (p=0.873) and hypothesis breadth (p=0.590) between control and experimental participants. This shows that the year in program did not have an effect on the results of the study.

CHAPTER 5: DISCUSSION

> "As [students] continue to work with more problems, especially in the problem-oriented learning system..., they accumulate more information that enriches their usable knowledge or experience base. This, in turn, will permit more numerous and effective hypotheses to be produced in the future. This is the objective of problem-based learning. It provides the [clinician] with associations in his long-term memory that will facilitate the generation of immediate hypotheses, making him more effective in his reasoning (Barrows & Tamblyn, 1980, p. 47)."

While the conventional hypothetico-deductive approach to reasoning in nursing is focused on the process of hypothesis retrieval and testing, abductive reasoning is focused on the process of hypothesis linking and creation (Mirza et al., 2014a; Raholm, 2010a; Vertue & Haig, 2008). Proponents of abductive reasoning such as Eriksson and Lindström (1997), and Ward and Haig (1997) regard hypothetico-deductive reasoning as a narrow approach to hypothesis generation which minimizes detection of underlying relationships in the data. As a result, these scholars propose abductive reasoning as an alternative approach which offers a broader approach to hypothesis generation. By building on the process of abductive reasoning outlined by Vertue and Haig, this study explored whether or not abductive reasoning training can improve hypothesis generation abilities (accuracy, expertise and breadth) of novice baccalaureate nursing students when compared to the conventional hypothetico-deductive reasoning which is promoted in baccalaureate nursing education (Rideout & Carpio, 2001; Tanner, 2006; Wong & Chung, 2002).

To answer this question, a quasi-experiment with a control group and a follow-up was conducted with 64 participants (29 control and 35 experimental) enrolled in the School of Nursing at McMaster University where the problem-based learning pedagogy is used. During the training which was provided to both groups, control participants explored a care scenario in a 40-minute problem-based learning tutorial format while experimental participants explored a care scenario as part of a 40-minute abductive reasoning training session. In the abductive reasoning training, participants learned the steps involved in the abductive reasoning process as proposed by Vertue and Haig (2008). Participants: (a) detected issues (phenomena); (b) proposed possible causes (causal mechanisms) of detected phenomena; (c) constructed a visual illustration (causal model) by interconnecting phenomena and causal mechanisms with one another and new ideas; and (d) synthesized the causal model to formulate hypotheses to explain the care scenario.

The three main outcome variables which were examined in this study were the three main abilities of hypothesis generation based on the work of Ingram et al. (1998). These included:

1. *Hypothesis accuracy:* Number of correct hypotheses (that matched the expert list) as a percentage of the total number of hypotheses generated by the participant for a particular care scenario;
2. *Hypothesis expertise:* Number of correct hypotheses (that matched the expert list) as a percentage of the total number of hypotheses on the expert list for a particular care scenario; and

3. *Hypothesis breadth:* Number of biopsychosocial (biological, psychological, social) categories which were addressed by the hypotheses which participants generated for a particular care scenario.

While control and experimental groups were similar in their demographic characteristics and baseline performance, statistically significant improvements were noted in experimental participants on all three variables at one-week follow-up. These outcomes show the usefulness of abductive reasoning training in enhancing the development of hypothesis generation abilities among first and second year baccalaureate nursing students. The remaining of this chapter will discuss key issues in nursing education and practice which abductive reasoning could address based on the findings of this research study.

Hypothesis Accuracy

The mean hypothesis accuracy score (ratio between correct and total hypotheses) for experimental participants was similar at baseline and at one-week follow-up. However, these participants generated almost three times more total and correct hypotheses at follow-up when compared to their baseline performance. This indicates that although certain participants may generate more correct hypotheses, their hypothesis accuracy score will remain 0.2 (20%). This could be why the mean hypothesis accuracy score of experimental participants at follow-up was similar to their baseline score.

While there was a nearly threefold increase in the number of correct and total hypotheses generated by experimental participants, no increase in total and correct number of hypotheses was observed among control participants at one-week follow-up.

Rather, the hypothesis accuracy score of control participants decreased at follow-up when compared to their baseline pre-test score. This is because the mean number of total hypotheses generated by control participants increased slightly (3.14 at baseline versus 3.84 at follow-up) while the mean number of correct hypotheses decreased slightly (0.29 at baseline versus 0.21 at follow-up), causing a wider gap between total and correct number of hypotheses. Due to this gap, a decrease in hypothesis accuracy was observed in the control group at one-week follow-up.

Although hypothesis accuracy was one of the main study outcomes, it may not be of great importance for novice first and second-year baccalaureate nursing students as much as it may be for senior third and fourth year baccalaureate nursing students. This is because novice nursing students possess limited nursing knowledge and experience, and may be limited in their ability to detect pertinent phenomena about a care situation. As a result, their ability to generate hypotheses may be prone to errors and poor hypothesis accuracy. Furthermore, the learning focus of novice nursing students is not on generating accurate hypotheses; rather, is on generating a broad range of hypotheses in order to identify and recognize several factors that influence the health of the person in a care situation (McMaster University BScN Handbook, 2012).

Hypothesis accuracy is more crucial for senior nursing students who, by graduation, are expected to provide errorless and competent care after they have correctly recognized and integrated a person's health issues (Del Bueno, 1994; 2005). These expectations are also congruent with the expectations of the School of Nursing at McMaster University, where senior nursing students are expected to possess increased knowledge and experience which is necessary for the correct detection and interpretation of pertinent client data and accurate generation of hypotheses (McMaster University BScN Handbook, 2012; Rideout & Carpio, 2001).

Hypothesis Expertise

Presented in Figure 5.1 is a pictorial representation of the ideas of a group of control participants as they attempted to explain their understanding of a care scenario during the 40-minute group discussion. The pictorial represents possible issues but does not show how various issues may be associated with one another. In contrast, presented in Figure 5.2 is a pictorial representation (i.e., causal model) of the ideas of a group of experimental participants as they attempted to explain their understanding of the same care scenario during the 40-minute abductive reasoning training. Different from the control group, the experimental group generated a series of issues and analyzed them by demonstrating and explaining various associations between them and their possible causal mechanisms.

While several issues were important, salient issues were determined as those which drew increased number of connections with other issues. In the pictorial (Figure 5.2), these are represented with a square box around them. These salient issues were similar to the ideal list of responses generated by the experts. Hypotheses connecting salient issues were also similar to the ideal list of responses generated by the experts. This shows that although nursing students may generate a variety of issues when they encounter a care situation, the abductive reasoning process could allow them to recognize which of these issues are salient based on the various connections that may exist between them. This is particularly important because educators are being encouraged to promote nursing students' abilities in prioritizing data and detecting salient issues (Benner et al., 2010; del Bueno, 1994, 2005).

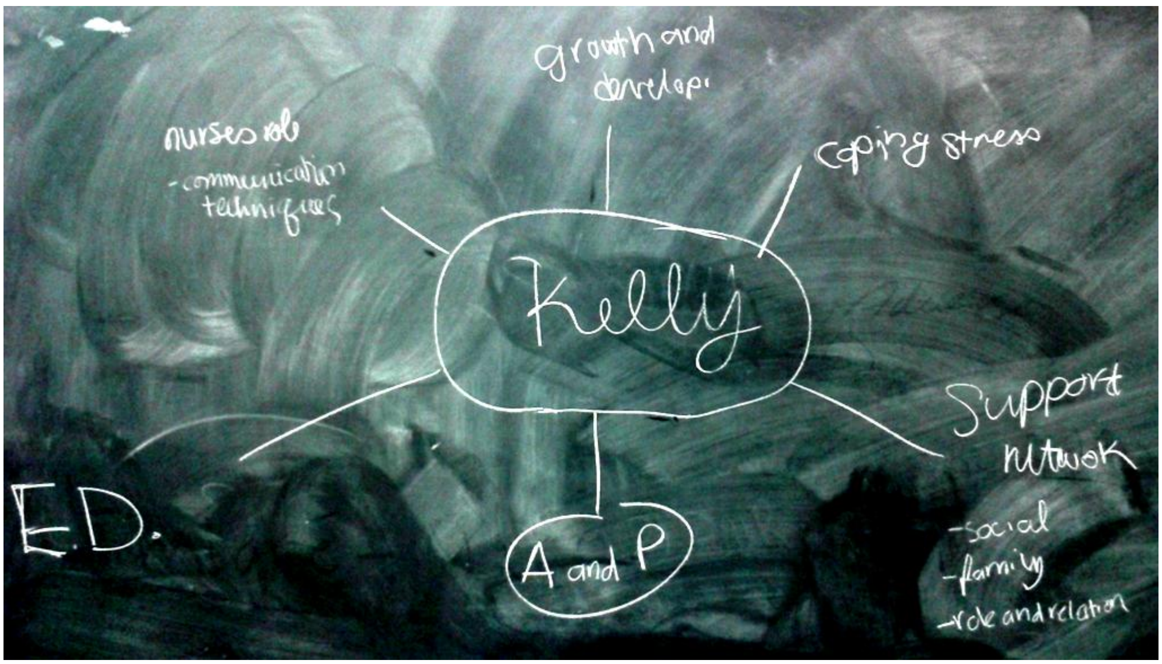


Figure 5.1. Pictorial representation of a control group's ideas about a care scenario during the group discussion.

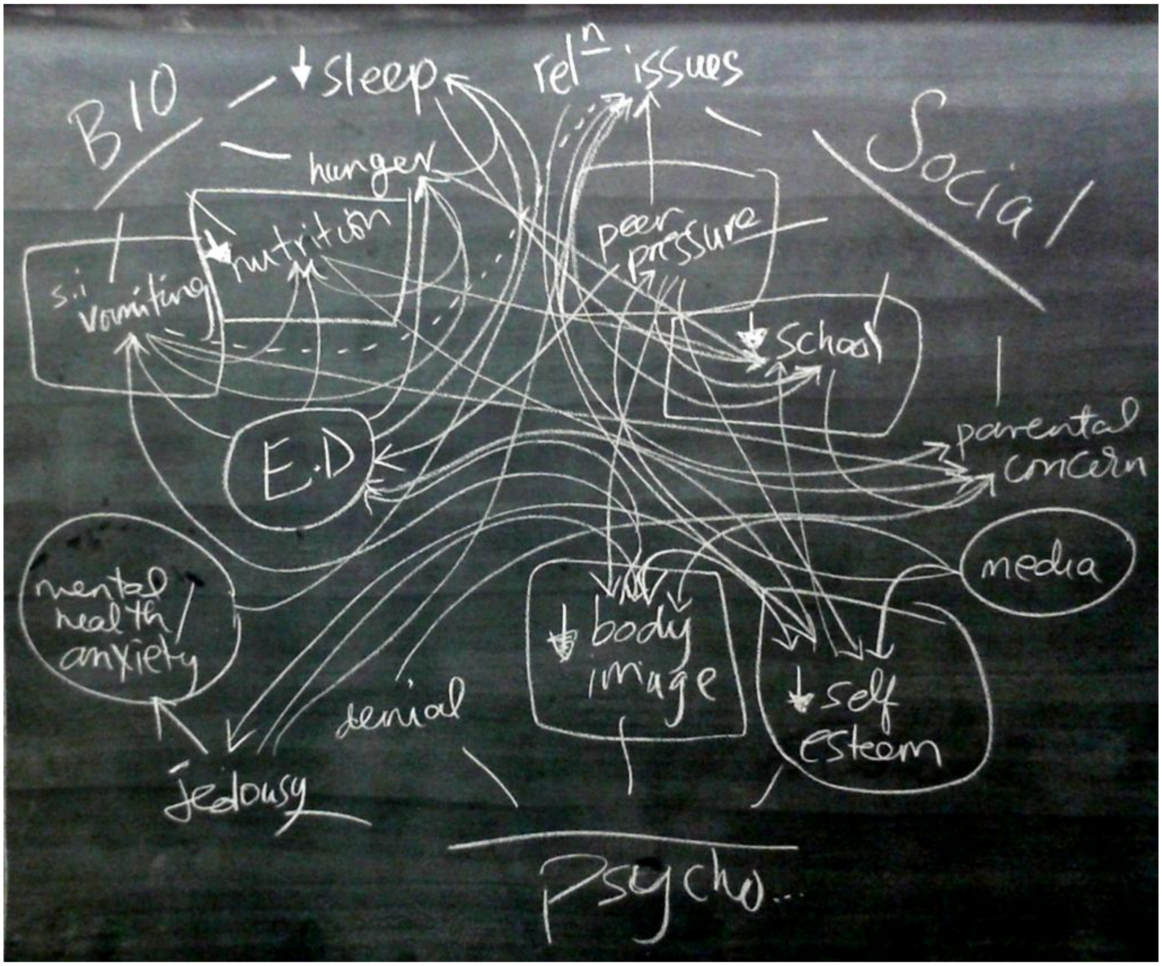


Figure 5.2. Causal model based on abductive reasoning. This represents an experimental group's ideas about a care scenario as they apply principles of abductive reasoning during the training session.

Since salient points discovered in the study by experimental participants were similar to the expert list of responses, salience was noted to be linked with expertise. While an expert has traditionally been known as a person with a lot of knowledge and skills (Benner, 1984), the modern expert has been described as a *relational learner* who "knows how to access knowledge efficiently and judiciously and who can form conceptual links between seemingly unrelated areas" (Fraser & Greenhalgh, 2001, p. 800). Hence, a person who can draw on previous knowledge and use it to form links between different ideas in a new care situation can be considered as having expertise. In this study, this was demonstrated by experimental participants who not only developed the ability to interlink various phenomena and causal mechanisms, but also transferred this ability to a new care scenario which was presented at one-week follow-up.

Steps of abductive reasoning explored in this study corresponded to Fraser and Greenhalgh's (2001) description of expertise because they allowed participants to integrate their existing knowledge, identify conceptual links between issues, and discover salient points. This may be why experimental participants who received abductive reasoning training were able to recognize more salient points of a care scenario and performed significantly better on hypothesis expertise when compared to control participants. Enabling nursing students to detect and interpret client cues, integrate previous knowledge, and analyze and synthesize data to determine salient points and the relationships between them has been a challenge for educators who have highlighted the need for active learning strategies that target these areas (Benner et al., 2010; del Bueno, 1994, 2005; Purling & King, 2012). This study demonstrates that abductive reasoning training may be a useful teaching and learning strategy in addressing this need in nursing education.

Hypothesis Breadth

While all study participants had knowledge of ways of knowing and the biopsychosocial framework, breadth categories in the study were limited to the biopsychosocial framework by the researcher. This simplified abductive reasoning training and the development of the causal model. Since abductive reasoning training was delivered to experimental participants in small groups, the pictorial depicted in Figure 5.2 is an example of a causal model that was constructed by one small group of experimental participants as part of the abductive reasoning training. However, after the group training, experimental participants constructed individual causal models at immediate post-test and one-week follow-up. This indicated that the construction of causal models is not solely a group activity; rather, it can be completed individually. Control participants, on the other hand, did not create causal models as it is not a strategy that is used in problem-based learning tutorials. However, a group of control participants did create a diagram of their ideas which is presented in Figure 5.1.

When compared to the diagram of control participants' ideas (Figure 5.1), the abductive causal model created by experimental participants (Figure 5.2) shows their ability to think more broadly by making connections between issues from various biopsychosocial categories. By integrating various health-related issues, experimental participants demonstrated an increased conceptual understanding about the breadth of the complexity involved in a particular care scenario. Not only this, they also demonstrated depth in their thinking, where depth is related to the complexities and multiple interrelationships which statements (i.e., hypotheses) contain (Paul & Elder, 2014). While the richness of the causal model was an unexpected study finding, it could explain why experimental participants performed significantly better than control participants in hypothesis breadth at immediate post-test and one-week follow-up.

Since care scenarios were written clearly to ensure participants were able to detect a broad range of issues, both control and treatment participants detected a variety of biopsychosocial issues. At all study time points, both groups consistently scored similarly on issue breadth. This emphasized that nursing education in the BScN program at McMaster University promotes broad-level thinking among its students and encourages them to think about the different aspects of a person's health. However, this broad-level thinking was limited to issue formation only. In terms of linking different aspects of health with one another through broad-level hypotheses, significant improvements were observed only in the experimental group which received the abductive reasoning training.

Effectiveness of Using Abductive Reasoning as a Teaching-Learning Strategy

Connectionism and Complexity

Newell and Simon's (1972) theory of information processing influenced hypothetico-deductive reasoning within medicine (Elstein et al., 1978) and problem-based learning (Barrows & Tamblyn, 1980). Their theory was based on an information processing system that operated in a serial or sequential manner, which means that an individual could only perform limited number of information processing tasks at a given time. This is why clinicians were known for having the ability to entertain a limited number (5 ± 2) of hypotheses when exploring a care situation (Barrows & Tamblyn; Elstein et al.). Since the information-processing model proposed by earlier researchers was too limiting, researchers in the 1980s adopted connectionism as a new way of information processing (Chandrasekaran, Goel & Allemang, 1988).

Introduced in the field of artificial intelligence by McClelland, Rumelhart, and the PDP (Parallel Distributed Processing) Research Group (1986), connectionism is founded on the idea that the brain as a processing system consists of patterns of units known as nodes which are connected with one another by several linkages. These connections form neural networks which resemble synaptic structures within the brain. Different from the previous information processing system that allowed the processing of only a limited number of tasks at any given time (Newell & Simon, 1972), connectionist approaches to information processing were viewed as being more holistic in nature, and could operate in a parallel manner, which meant that several tasks could be processed at once at any given point in time (Chandrasekaran et al., 1988).

As a parallel approach to information processing, connectionist models in the field of artificial intelligence attempt to explain human intellectual abilities by using computer modeling through artificial neural networks (McGonigle & Mastrian, 2012). While connectionist models have been linked to abductive reasoning in the field of artificial intelligence (Ajjanagadde, 1993; Johnson, Zhang & Wang, 1997; Roth, 1996; Wang, Johnson, & Zhang, 2006), the relationship between the two has not been discussed in the context of clinical reasoning in nursing or in the work of Vertue and Haig (2008).

Therefore, it is important to highlight the potential link between connectionist models and the causal models within abductive reasoning.

The causal model depicted by the experimental group in Figure 5.2 resembles the complex, holistic, and integrated nature of connectionist models and the brain in general. It also shows the formation of several connections between numerous issues at the same time (i.e., parallel processing). This model also prevents information overload which could occur in the traditional information processing system due to its limited processing ability. The resemblance of the visual causal model technique of parallel processing may be one reason why experimental participants generated issues and hypotheses which surpassed the theorized capacity estimates of the number of issues and hypotheses which are generally expected from a person based on the traditional theory of the human information processing system by Newell and Simon (1972).

In addition to connectionism and neural networks, the causal model also resembles complex systems models. As a new approach to science, the study of complex systems examines how relationships between various components of a system give rise to the collective behaviors of the system as it interacts with the environment (Bar-Yam, 2009). Building on this, the present study exposed experimental participants to the idea of the person as a complex system where various biopsychosocial aspects of health are linked to one another and allow the formation of broad hypotheses that best-explain how one person's health may be influenced by various interrelated factors. This is congruent with abductive reasoning and its focus on allowing individuals to comprehend a broad and complex reality (Eriksson & Lindström, 1997; Raholm, 2010a, 2010b; Vertue &

Haig, 2008), rather than the reductionist perspective of empiricism from which hypothetico-deductive reasoning emerged (Monti & Tingen, 1999).

Many experimental participants verbally expressed that the care scenarios used in this study became increasingly difficult from the first to the last. However, their scores on the three study outcomes (hypothesis accuracy, expertise and breadth) became significantly higher than control participants as both groups progressed through the care scenarios. While content experts had rated all care scenarios as similar in level of difficulty, experimental participants could have found the latter scenarios as more difficult due to the complexity lens which they could have developed through abductive reasoning training. This was demonstrated in participants' breadth scores. While both groups were able to detect broad-level issues in each care scenario, only the experimental participants generated more broad-level hypotheses after the abductive reasoning training. This could be due to their beginning ability to grasp the complex nature of the case scenarios.

Information Processing Capacity

Since the environment often provides more input than the human information processing system can accept and process, an individual copes with this overload by processing a limited number of tasks through filtering information. This filtering is based on the individual's prior knowledge and experience, and certain time pressures (Oke, 2009). In the context of hypothesis generation, the hypothesis capacity in the clinical reasoning process (hypothetico-deductive reasoning) according to Barrows and Tamblyn (1980) is usually not less than 3 and not more than 7 (i.e., 5 ± 2). This same capacity is estimated at approximately 4 hypotheses by Elstein et al. (1978). However, hypothesis generation under time pressure is known to result in the generation of fewer hypotheses (Dougherty & Hunter, 2003; Thomas, Dougherty, Sprenger & Harbison, 2008), sometimes even a single hypothesis (Flin, Slaven, & Stewart, 1996; Klein, 1993).

It is noteworthy that Elstein et al. (1978) and Barrows and Tamblyn (1980) refer to hypotheses as early medical diagnoses (i.e., one-word disease descriptors). However, these hypotheses are considered as descriptive hypotheses (referred to as *issues* in this study) involving one variable, as compared to relational hypotheses (referred to as *hypotheses* in this study) which involve a relationship between two variables (Sumathi & Saravanavel, 2008). This distinction between descriptive and explanatory hypotheses is also vague in the *Person-Based Learning within Problem-Based Learning Model* used by the School of Nursing at McMaster University (McMaster University BScN Handbook, 2012), which uses the two terms interchangeably (i.e., "issues/hypotheses"). Therefore, while the 5 ± 2 hypothesis capacity estimate could be used as a guide to examine descriptive hypotheses (i.e., issue generation) of all study participants, no specific guide can be used as a guide for relational hypotheses (i.e., hypothesis generation).

In the present study, participants were required to generate issues and hypotheses. The mean number of issues generated by both control and experimental groups at baseline (i.e., 6.5 and 6.3 respectively) was within the theorized capacity estimate of 3 to 7 issues. Control participants continued to generate issues close to this capacity estimate at both the immediate post-test (7.5) and at one-week follow-up (7.6). However, after the abductive reasoning training, the mean number of issues generated by experimental participants surpassed the theorized capacity estimate of 3 to 7. Experimental participants generated 11.8 issues at immediate post-test and 11.4 issues at one-week follow-up. While the focus of this study was hypothesis accuracy, expertise and breadth, the significant increase in the mean number of total issues generated by the experimental group after the abductive reasoning training was an unexpected study finding.

A similar pattern was observed when participants generated hypotheses which explained a relationship between at least two issues. The mean numbers of hypotheses generated by the control group at baseline pre-test was 3.1, and 4.1 at immediate post-test and 3.8 at one-week follow-up. The mean number of hypotheses generated by the experimental participants at baseline was lower than the control participants (i.e., 2.3 versus 3.1). However, after receiving the abductive reasoning intervention, experimental participants generated 6.2 hypotheses at immediate post-test and 6.1 hypotheses at one-week follow-up. This number of hypotheses was significantly higher than the hypotheses generated by control participants at immediate post-test (4.1) and one-week follow-up (3.8). This was also an unexpected study finding.

Since the causal models took time away from the allotted amount of total time (i.e., 15 minutes) to generate issues and hypotheses, experimental participants were under more pressure than control participants to generate hypotheses. Due to this time pressure, experimental participants would be expected to generate fewer hypotheses as proposed by the literature which states that the number of hypotheses decreases under pressure (Dougherty & Hunter, 2003; Flin, Slaven, & Stewart, 1996; Klein, 1993; Thomas, Dougherty, Sprenger & Harbison, 2008). However, study findings revealed that despite time pressure, experimental participants who received abductive reasoning training generated more issues and hypotheses than control participants and some of the hypothesis capacity estimates proposed by scholars (Barrows & Tamblyn, 1980; Elstein et al., 1978).

The significant improvement in the experimental group's mean issue and hypothesis scores after receiving the abductive reasoning training may be attributed to the training itself. Abductive reasoning allowed experimental participants to form connections between multiple phenomena and causal mechanisms that affect a particular client care situation. The formation of connections between various factors relates back to connectionism and neural networks, which allow the simultaneous processing of multiple data (Chandrasekaran et al., 1988). The network of relationships portrayed in a causal model (Figure 5.2) also helps an individual understand the complex nature of the care situation. This is the main purpose of abductive reasoning – i.e., to help an individual recognize meaningful underlying patterns of selected phenomena so hypotheses which explain a complex reality can be created (Eriksson & Lindström, 1997; Raholm, 2010a, 2010b).

Development of Capability

According to Oke (2009), memory storage consists of frames of reference which are patterns of an individual's understanding of how ideas are related to one another within a particular context. These frames of reference are based on prior knowledge and experience, and their effective use is known to be a characteristic of expertise (Oke). However, errors in inference may arise if the detection of pertinent data is insufficient or inconsistent with established frames of reference due to limits in human memory (Oke). In the context of information processing, frames of reference allow individuals to recognize known patterns in newer situations (Oke). This process is referred to as pattern recognition or intuition (Benner, 1984; Tanner, 2006).

While pattern recognition is a common method used by clinicians for detecting patterns in data (Norman et al., 1994), Fraser and Greenhalgh (2001) discuss another transformational process known as capability, where learners adapt and tune existing competencies to newer situations. According to Fraser and Greenhalgh, this is especially useful when dealing with complex situations where pre-existing patterns may not be able to appropriately inform all aspects of a newer complex situation. The development of capability could enable learners to work effectively in unfamiliar situations that require them to apply knowledge in ways that are beyond the scope of their textbooks (Fraser & Greenhalgh), and potentially even beyond established frames of reference that exist in their memory.

As this study shows, all participants integrated previous frames of reference to understand the care scenario at baseline. To do this, they used various strategies outlined in Table 1.1. First, participants made note of cues (i.e., noticing, cue acquisition) mentally or by highlighting the scenario text. They then formulated a list of issues based on pre-existing frames of reference from their knowledge base (i.e., retrieval of hypotheses, pattern recognition). By re-reading the text, they searched for more cues that would further support their list of issues (i.e., cue interpretation) before deciding on one or more frames of reference as the main explanatory hypotheses which could explain some or all of the issues (i.e., confirmation of hypothesis). While control participants repeated this process at immediate post-test and at one-week follow-up, experimental participants used a different approach.

After the abductive reasoning training, experimental participants not only used pre-existing frames of reference to detect phenomena (issues), they also demonstrated capability by adapting and tuning their frames of reference to accommodate the dynamic patterns of the new care scenario. This was done by inferring causal mechanisms, interconnecting them with various phenomena through the development of a broad-level casual model, and then synthesizing this understanding through the creation of relational hypotheses which were specific to the issues of the person in the care situation. While experimental participants generated hypotheses that focused on the salient biopsychosocial issues presented in the care scenario, control participants generated hypotheses that were general and slightly unrelated (Appendix 12 & 13).

Some control participants (approximately 20%) used their hypotheses from the baseline scenario to also explain the new scenario at immediate post-test. Since the time between baseline pre-test and immediate post-test was approximately 40 minutes, control participants may have experienced mental overload when trying to process data from two care situations within a short period of time. Hence, when presented with a new scenario (involving the same client population – i.e., female adolescents), control participants may have begun making assumptions and could have generated early hypotheses based on their recent interaction with the baseline scenario. This may reflect a characteristic of hypothetico-deductive reasoning. Experimental participants, however, generated scenario-specific hypotheses and did not use hypotheses from the baseline scenario to explain the new scenario. This could be due to the connectionist approach of abductive reasoning which may have enabled them to process multiple data within a short period of time without experiencing mental overload.

While control participants' reuse of previous hypotheses to explain a newer situation was an unexpected finding, it may portray one limitation of the sequential information processing system which can only process a limited number of commands at a given time as compared to the connectionist model where multiple tasks can be processed simultaneously. This finding may also portray a limitation of hypothetico-deductive reasoning where clinicians are known to generate early hypotheses based on pre-existing hypotheses, after which they search deliberately for additional data to support their early claims (Norman et al., 1999; Simmons et al., 2003). To overcome these limitations, Fraser and Greenhalgh's (2001) idea of developing capability can be useful because it not only allows learners to use existing competencies to understand newer situations; it also encourages a fresh lens instead of fitting situations into pre-existing frames of reference.

Novice versus Expert

In the past, experts have been known as persons with a lot of knowledge and skills gained through year of experience (Benner, 1984). However, the modern expert is described as a person who can interlink unrelated ideas (Fraser & Greenhalgh, 2001). In this study, abductive reasoning was found to be useful in building capability to enhance expertise of baccalaureate nursing students. Several observations in this study show that experimental participants were beginning to identify more salient points which matched the salient points identified by experts. This supported the procedural definition of expertise which describes hypothesis expertise as the total number of correct hypotheses divided by the number of hypotheses identified by the experts (Ingram et al., 1998).

According to Glaser and Chi (1988), experts aim to understand a problem while novices try to immediately apply equations to resolve the problem. This was evident when control participants incorrectly re-used some of their earlier hypotheses in a newer situation, while experimental participants, after abductive reasoning training, began making attempts to understand the newer situation and its complex parts before generating hypotheses. Glaser and Chi also indicate that experts understand situations better than novices because they build mental models which structure their knowledge and allow them to infer associations that help them understand the presented situation. Congruent with this expert attribute, experimental participants in this study demonstrated an improved ability to develop causal models to help them apply, organize, and interlink their knowledge so they can better-understand newer and unfamiliar care situations.

While control participants did not develop causal models, a group of control participants did, however, develop a pictorial of their ideas (Figure 5.1). This pictorial shows less depth and breadth than a causal model developed by a group of experimental participants (Figure 5.2). The pictorial developed by control participants portrays a superficial understanding of the care situation while the causal model developed by experimental participants exhibits more depth (i.e., a more organized and interrelated understanding of the care situation). The depth and integration portrayed in the causal model which allows for the recognition of salience, show that abductive reasoning could assist novices to begin to think more like experts. This finding is similar to early experiments in the field of physics (Chi, Feltovich & Glaser, 1981) and programming (Weiser & Shertz, 1983) which showed that experts are more likely to use more meaningful and methodical approaches when explaining problems while novices are more likely to use superficial approaches to explain problems.

Novice nurses, including nursing students, are known to use hypothesis testing (i.e., hypothetico-deductive reasoning) as compared to expert nurses who come up with conclusions through intuition or pattern recognition without having to test their hypotheses (Benner, 1984; Tanner, 2006). Nursing scholars believe that nursing students use hypothetico-deductive reasoning because it can be taught (Harbison, 1991; Rideout & Carpio, 2001). Contrary to this, it is believed that intuitive judgment, which is used by experts, cannot be taught because it develops only with knowledge and experience-based expertise (Benner). However, this study shows that abductive reasoning can be taught and could be useful in building capability that could result in enhanced expertise of baccalaureate nursing students. By enhancing expertise through abductive reasoning, educators may be able to help students build capability to think more like experts (i.e., an increased recognition of salience), and facilitate their journey from novice to expert. The development of such expertise would mean that individuals will be able to engage in relational thinking by forming conceptual links between unrelated ideas in order to understand unfamiliar situations.

Implications

Baccalaureate Nursing Education

Success in nursing practice is constantly being redefined in terms of skills required to effectively manage healthcare situations with increasing acuity and complexity (Li & Kenward, 2006; Purling & King, 2012). Abductive reasoning could be useful in the classroom setting as a scaffolding strategy which aims to engage students with varying interests. For example, if one student is interested in physiological aspects of a care situation, another student may be interested in social aspects, and so forth. Together, students within a group can construct a complex causal model which depicts the care situation as a complex system where all parts of the model are linked with one another. This can not only enrich the quality of discussion, it can also make learning more meaningful for students who may have increased opportunities to contribute to each step of the abductive reasoning process.

With further use of abductive reasoning in baccalaureate nursing education, educators may be able to promote a broader approach to understanding care situations. Continuous engagement in abductive reasoning and the construction of causal models based on several care situations could give way to the development of strong and well-organized associations in the long-term memory (i.e., frames of reference), which may facilitate the generation of immediate hypotheses in future care situations (Barrows & Tamblyn, 1980). It is, therefore, important to focus on performance rather than competence alone (Miller, 1990). By doing so, educators could educate for capability (Fraser & Greenhalgh, 2001), which will allow nursing students to adapt to changing care situations within the classroom, generate new knowledge, and improve their performance within both the classroom and the clinical setting on an ongoing basis.

Although most new nursing graduates pass competency examinations, many struggle to deal with complex situations in their clinical practice (del Bueno, 1994, 2005). Since performance is the stage after competence (Miller, 1990), the positive effects of abductive reasoning on performance, as demonstrated in this study, make it a useful teaching and learning strategy to promote within clinical nursing education as well. This strategy could be used within post-conferences to allow nursing students to explore a care situation in more depth through the analysis and synthesis of various factors that affect the client's health. However, the process of abductive reasoning would need to be initiated by the educator for the first few sessions until the learners can initiate it on their own. For this reason, educators would need to be trained on how to engage nursing students in abductive reasoning. A long-term consequence of this would be that by senior years, nursing students may become so proficient at developing causal models that they may be able to create them mentally which would lead to more efficient nursing practice.

During clinical nursing education, abductive reasoning could be useful in helping nursing students understand health situations of varying complexity. Since about 75% of middle aged and older adults have at least two co-existing health conditions (Schoenberg, Bardach, Manchikanti & Goodenow, 2011), abductive reasoning could be particularly useful in helping nursing students understand and explain the complexity associated with multimorbidity among older clients. However, abductive reasoning is not limited to older clients. It could be used with a variety of client populations and in varying settings where complex situations are found. Lunney (2008) highlights a critical need to address the accuracy of nurses' hypotheses through training. She further encourages educators to teach nursing students ways to generate more accurate relational hypotheses. Therefore, training nurses in abductive reasoning could be useful in enhancing the accuracy of their hypotheses. This could be particularly useful at the School of Nursing at McMaster University where no formal training on hypothesis generation is provided to baccalaureate nursing students (Jewiss, T., First Year BScN Chair McMaster University, personal communication, 2012).

The use of abductive reasoning as a teaching and learning strategy could also allow educators to increase the depth and breadth of conceptual thinking among novice nursing students. As a result, this could lead to more meaningful learning for students and an increased capacity to understand complex health situations earlier in the baccalaureate curriculum. Early exposure to strategies which enhance understanding of complex health situations could ensure that when novice nursing students enter senior years, they will begin to view a client's health conditions as integrated within one body rather than separate and independent entities. This may also lead to a more holistic and person-based (rather than problem-based) approach to care which fits with the aim of the *Person-Based Learning within Problem-Based Learning Model* used by the School of Nursing at McMaster University (McMaster University BScN Handbook, 2012).

The School of Nursing at McMaster University also utilizes Tanner's (2006) C*linical Judgment Model*, which describes how experienced nurses think through complex care situations, and aims to teach nursing students this expert way of thinking.

Since this model is similar to the nursing process, it presents another creative way to teach hypothetico-deductive reasoning. It does not, however, teach intuition or a way to develop expertise because, according to Benner (1984) intuitive judgment cannot be taught and must be developed through expertise. The first stage of abductive reasoning corresponds with the first two stages of the Tanner's model (i.e., noticing and interpreting). While Tanner's *interpreting* stage uses analytic, intuitive, and narrative reasoning patterns; missing in this approach is the capability to create a synthesis within clinical reasoning which abductive reasoning promotes (Eriksson & Lindström, 1997; Raholm, 2010a; Vertue & Haig, 2008). As shown in this study, the synthesis which abductive reasoning promotes could enhance hypothesis expertise and capability of nursing students. This may also contribute to an early development of intuitive judgment which is an attribute of nurses with more expertise.

Practice of Licensed Nurses

Since the healthcare system demands nurse educators to prepare nursing graduates who can adapt to the changing needs of clients (Wolff, Regan, Pesut & Black, 2010), the development of capability among new graduates is crucial. As an observed outcome of abductive reasoning, capability may enable nursing graduates to use their existing understanding about complex health situations and adapt it to newer complex health situations. This may lead to more efficient delivery of nursing care. There are certain entry-to-practice competencies which abductive reasoning may promote among new graduates. Based on the *Competencies for entry-level Registered Nurse practice* (CNO, 2014, p.7), these are:

1. Analyzes and interprets data obtained in client assessments to draw conclusions about client health status (*Competent Application of Knowledge: Ongoing Comprehensive Assessment, competency #40).*
2. Anticipates potential health problems or issues for clients and their consequences and initiates appropriate planning (*Competent Application of Knowledge: Collaborating with Clients to Develop Health Care Plans, competency #49).*

Since abductive reasoning enhanced nursing students' expertise in this study, it may also be useful in clinical practice of novice nurses (including new graduates). Novice nurses engaged in abductive reasoning may find it useful in gaining more expertise in the specialty area within which they practice. New graduate transition programs could train nurse mentors on how to engage their mentees in abductive reasoning so they can promote learning about the complexity involved in care situations. This will also enable them to exercise the capability which they would have developed from their previous exposure to abductive reasoning in baccalaureate nursing education.

Abductive reasoning can be useful within the clinical context where nurses are required to collaborate with other health disciplines and think about a client's health from a holistic point of view. The breadth in thinking that abductive reasoning promotes could assist nurses in understanding and explaining various interrelated factors that affect clients' health. However, this holistic lens needs to be enhanced through training of novice nurses in the clinical setting. To accomplish this, abductive reasoning workshops, similar to the 40-minute training session used in this study, could be introduced during nursing orientation or even as part of the new nursing graduates' transition programs. Workshops or re-training sessions could also be offered to experienced nurses and, potentially, to practitioners of other health disciplines so they can take advantage of the broad approach to thinking which abductive reasoning offers.

According to Lunney (2008), there is a need to address the accuracy of relational hypotheses of nurses. Since correct interventions and desired client outcomes depend highly on accurate clinical judgments about complex care situations (Levin, Lunney & Krainovich-Miller, 2004), abductive reasoning may be useful for nursing practice. Abductive reasoning promotes a capability which could promote hypothesis generation abilities such as hypothesis accuracy, broad-level thinking and discovery of salient aspects of a care situation. Improved hypothesis generation abilities may be useful in planning nursing interventions which are more person-centered and take into consideration multiple factors which affect a client's health. This could further lead to enhanced quality of care that nurses provide to clients, especially those who live with complex health challenges.

As previously discussed, the key reasoning stage in Tanner's (2006) model is interpreting, where the nurse may use one or a combination of analytic, intuitive, and narrative modes of reasoning. While this variation in options provides flexibility in thinking, this same variety may create issues for effective nursing practice. In her analysis of 20 studies, Lunney (2001) discussed that nurses' interpretation of the same data varies widely. Since the time of Lunney's analysis, several other studies have also supported this finding (Lunney, 2008). This gap could be addressed by abductive reasoning which allows individuals to generate and analyze a variety of ideas. However, with the development of the causal model, they can begin to discover key salient points of care situations. As shown in this study, through the synthesis which abductive reasoning promotes, individuals can begin to make similar interpretations (i.e., common salient points) which match more closely to the list of salient points generated by experts.

Future Directions

In this study, experimental participants did not receive any feedback on their newly developed abductive reasoning skills when they rehearsed them at immediate post-test. Should this study be repeated, participants may benefit from receiving feedback on their causal models. This will ensure participants have understood the abductive reasoning process correctly, may result in improved causal model development, and could contribute differently to the study results. Since this study involved single-session training, there is a need for future research that explores the influence of multi-session abductive reasoning training on hypothesis generation abilities of baccalaureate nursing students. With two to three study sessions committed to the rehearsal of abductive reasoning skills prior to follow-up, participants could have sufficient opportunities to practice and obtain feedback on their skills in abductive reasoning and causal model development. Ongoing training of this kind may better influence scores than a one-time training and application session.

In future research involving abductive reasoning, varying care scenarios which differ in theme and complexity could also be used to determine how participants in both control and experimental groups fare on study outcomes (i.e., hypothesis accuracy, expertise and breadth). Variation in care scenarios may present a more realistic depiction of how abductive reasoning can influence hypothesis generation abilities of nursing students in different complex client populations, particularly in the area of complex chronic care. Furthermore, future studies similar to this study could also include senior students in order to examine how abductive reasoning as a scaffolding strategy could potentially help third and fourth-year students think about complexity in care situations. In addition to this, qualitative research could also be conducted using focus groups with the participants in each training session. This could allow nurse educators to understand the usefulness and practicality of abductive reasoning in helping nursing students understand complex health situations.

While abductive reasoning can allow nurses and nursing students to understand the complexity of clients' health-related issues, it may also be useful in helping them understand the complexity involved in healthcare organizations and the healthcare system. This is outlined by Ebright, Patterson, Chalko and Render (2003) who examined various work-related complexities which impede nurses' ability to efficiently provide care. These included missing supplies, interruptions, geography of assignments, communication gaps, etc. This is why in addition to educating students about diseases, procedures and clinical reasoning, nursing curricula must also focus on processes around the management of workload complexities. The complexity lens which abductive reasoning promotes could be used to examine and understand how various workplace-related complexities are interrelated and how they may be affecting nurses' performance within an organizational context.

Furthermore, abductive reasoning could also be used to examine and understand how broader healthcare system issues are interrelated and how gaps in the system can inform future system-wide practice changes. Recently, the work of Bayliss et al. (2014) calls for a more integrated healthcare system by moving the healthcare system away from the reductionist approach of “What is the matter” to a system which values broader contextual factors and asks “What matters?” Abductive reasoning, as a framework, could potentially be useful in examining such broader system-related phenomena so new knowledge can be generated through a more integrated healthcare system. This could lead to improved and more efficient care for clients as they transition from one aspect of the system to the next.

To move abductive reasoning from the classroom to the clinical setting, the first step may be to test abductive reasoning training with simulations. The testing of abductive reasoning with simulation could shed light on whether or not this form of reasoning can enhance learning in a more hands-on environment. Should results be positive, abductive reasoning could then be tested in the clinical setting with students. The usefulness of abductive reasoning could also be tested with nurses through training sessions that are similar to the ones delivered in this study. Not only will this determine if abductive reasoning can enhance nurses’ understanding of complex health situations at the point of care, it will also indicate the training’s practicality within a clinical setting. Such explorations could contribute immensely to the limited literature on abductive reasoning within nursing education.

Study Limitations

A major study limitation was that only first and second year baccalaureate nursing students who were enrolled in high school prior to entering the BScN program were recruited. Therefore, findings from this study may not be generalizable to mature students and third and fourth-year baccalaureate nursing students. Another limitation was that the study power for hypothesis accuracy (one of the three outcome variables) was 0.503, which is lower than the 0.8 standard used in research (Hulley et al., 2007). Based on power analysis, the study power for hypothesis accuracy would have been 0.8 if 126 participants were enrolled in the study instead of 64, or if the effect difference was 0.1 instead of 0.065 between control and experimental groups. One way to attain a higher effect difference could be through more than one abductive reasoning training session. This could also allow participants to develop a stronger foundation in abductive reasoning.

Conclusion

Since researchers encourage nurse educators to use active learning strategies that emphasize the application of knowledge, analysis and synthesis of data, detection of salient points, and the prioritization of care needs (Benner et al., 2010; del Bueno, 1994, 2005; Li & Kenward, 2006; Purling & King, 2012), abductive reasoning training offers a unique teaching and learning scaffolding approach to tackling these challenges in which nursing students can adapt and tune their existing knowledge to new situations, and analyze and synthesize data in order to discover salient points and the key relationships between them. As shown in this study, this process could positively impact students' hypothesis accuracy, expertise and breadth.

This study shows that abductive reasoning can allow nursing students to link multiple ideas together and formulate a complex diagram which displays the complex reality of a care situation. Such strategies could be explored further in the context of complex chronic disease to understand how they might inform nursing students' and nurses' understanding of complex disease. With the use of causal models, nursing students may be able to comprehend not only how each chronic disease affects the client; but rather, how several chronic diseases, along with various external factors, are linked with one another and how this complexity impacts the health of the client. This form of complex and relational thinking may better-prepare new nursing graduates to manage complex health situations when they enter the workforce.

Not only could abductive reasoning help new nursing graduates meet certain entry-to-practice competencies related to analysis and interpretation of data, it can also facilitate their movement from novice to expert by helping them develop reasoning capabilities are similar to those of experts. The holistic lens which abductive reasoning offers could also contribute to the clinical practice of nurses who are required to think more broadly due to the ever-increasing holistic approach to care. Due to its focus on discovering underlying structures (i.e., salient points) of a complex reality, abductive reasoning may also help nurses in generating common interpretations of client data instead of interpretations that vary widely.

Since this research is the first of its kind, more research is required to build a body of evidence which shows the impact of abductive reasoning on hypothesis generation abilities of baccalaureate nursing students. This will enable decision-makers to determine the usefulness of this approach to reasoning and its ability to inform nursing education, baccalaureate nursing curricula, and clinical practices of nursing students. As a pioneer in abductive reasoning research within the context of nursing education, this study is related to many other ideas such as complex systems theory, connectionism, capability building, and relational learning. Further development of knowledge in abductive reasoning may create opportunities for more specific research which explores connections of abductive reasoning with other ideas mentioned herein.

References

Ajjanagadde, V. (1993). Incorporating background knowledge and structured explananda in abductive reasoning: A framework. *IEEE Transactions on Systems, Man, and Cybernetics, 23*(3), 650-664.

Alligood, M.R. (2010). The nature of knowledge needed for nursing practice. In M.R. Alligood (Ed.), *Nursing theory: Utilization & application* (4th ed., pp. 3-16). Maryland Heights, MO: Mosby.

Allen, D. (1990). The curriculum revolution: Radical re-visioning of nursing education. *Journal of Nursing Education, 29*(7), 312-316.

Banning, M. (2008). A review of clinical decision making: Models and current research. *Journal of Clinical Nursing, 17*(2), 187-195.

Barrows, H.S., & Feltovich, P.J. (1987). The clinical reasoning process. *Medical Education, 21*(2), 86-91.

Barrows, H.S. & Tamblyn, R.M. (1980). *Problem-based learning: An approach to medical education.* New York, NY: Springer Publishing Company.

Bar-Yam, Y. (2009). General features of complex systems. In L.D. Kielayte (Ed.), *Knowledge management, organizational intelligence and learning, and complexity – Vol. I* (pp. 43-95). Paris, France: UNESCO-EOLSS.

Bayliss, E.A., Bonds, D.E., Boyd, C.M., Davis, M.M., Fink, B., Fox, M.H., ... Stange, K.C. (2014). Understanding the context of health for persons with multiple chronic conditions: Moving from what is the matter to what matters. *Annals of Family Medicine, 12*(3), 260-269.

Begg, C., Cho, M., Eastwood, S., Horton, R. Moher, D., Olkin, I., ... Stroup, D.F. (1996). Improving the quality of reporting of randomized controlled trials: The CONSORT statement. *Journal of the American Medical Association, 276*(8), 637-639.

Bell, D.S., Harless, C.E., Higa, J.K., Bjork, E.L., Bjork, R.A., Bazargan, M., & Mangione, C.M. (2008). Knowledge retention after an online tutorial: A randomized educational experiment among resident physicians. *Journal of General Internal Medicine, 23*(8), 1164-1171.

Benner, P. (1982). From novice to expert. *American Journal of Nursing, 82*, 402-407.

Benner, P. (1984). *From novice to expert: Excellence and power in clinical nursing practice.* Menlo Park, CA: Addison-Wesley.

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

Benner, P., & Tanner, C.A. (1987). Clinical judgment: How expert nurses use intuition. *American Journal of Nursing, 87*(1), 23-31.

Benner, P., Tanner, C., & Chesla, C. (2009). *Expertise in nursing practice: Caring, clinical judgement, and ethics* (2nd ed.). New York, NY: Sprinter Publishing Company, LLC.

Bevis, E. (1988). New directions for a new age. In National League for Nursing (Ed.), *Curriculum revolution: Mandate for change* (pp. 27-52). New York, NY: National League for Nursing.

Bevis, E., & Murray, J. (1990). The essence of the curriculum revolution: Emancipatory teaching. *Journal of Nursing Education, 27*(7), 326-331.

Black, B.P. (2011). The history and social context of nursing. In K.K. Chitty & B.P. Black (Eds), *Professional nursing: Concepts and challenges* (6th ed., pp. 29-59). Maryland Heights, MO: Saunders Elsevier.

Brody, T.A. (1993). A minimal ontology for scientific research. In L. de la Peña & P. E. Hodgson (Eds.), *The philosophy behind physics* (pp. 84-94). Berlin, Germany: Springer-Verlag.

Bruner, J.S. (1960). *The process of education.* Cambridge, MA: Harvard University Press.

Bruner, J.S. (1961). The act of discovery. *Harvard Educational Review, 31*(1), 21-32.

Buckingham, C.D., & Adams, A. (2000). Classifying clinical decision making: A unifying approach. *Journal of Advanced Nursing, 32*(4), 981-989.

Burns, N., & Grove, S.K. (2004). *The practice of nursing research: Conduct, critique, & utilization* (5th ed.). St. Louis, MO: Elsevier Saunders.

Carper, B.A. (1978). Fundamental patterns of knowing in nursing. *Advances in Nursing Science, 1*(1), 13-24.

Chambers, D., Thiekötter, A., & Chambers, L. (2013). Preparing student nurses for contemporary practice: The case for discovery learning. *Journal of Nursing Education and Practice, 3*(9), 106-113.

Chandrasekaran, B., Goel, A., & Allemang, D. (1988). Connectionism and information-processing abstractions. *AI Magazine, 9*(4), 24-34.

Chi, M.T.H., Feltovich, P.J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-125.

Chinn, P. (1989). Feminist pedagogy in nursing education. In National League for Nursing (Ed.), *Curriculum revolution: Reconceptualizing nursing education* (pp. 9-24). New York, NY: National League for Nursing.

Chitty, K.K. (2011). The education of nurses. In K.K. Chitty & B.P. Black (Eds), *Professional nursing: Concepts and challenges* (6th ed., pp. 146-170). Maryland Heights, MO: Saunders Elsevier.

Choudhary, R., Dullo, P., & Tandon, R.V. (2011). Gender differences in learning style preferences of first year medical students. *Pakistan Journal of Physiology, 7*(2), 42-45.

Chow, S.C., Shao, J., & Wang, H. (2008). *Sample size calculations in clinical research* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.

Cleophas, T.J., Zwinderman, A.H., Cleophas, T.F., & Cleophas, E.P. (2009). *Statistics applied to clinical trials* (4th ed.). Netherlands: Springer.

CNO (College of Nurses of Ontario). (2014). *Competencies for entry-level Registered Nurse practice.* Toronto, ON: Author.

Coburn, J. (1974). I see and am silent: A short history of nursing in Ontario. In J. Acton, P. Goldsmith, & B. Shepard (Eds.), *Women at work, Ontario 1850-1930* (pp.126-163). Toronto, ON: Canadian Women's Educational Press.

Coderre, S., Mandin, H., Harasym P.H., & Fick, G.H. (2003). Diagnostic reasoning strategies and diagnostic success. *Medical Education, 37*(8), 695-703.

Crooks, D., Lunyk-Child, O., Patterson, C., & LeGris, J. (2001). Facilitating self-directed learning. In E. Rideout (Ed.), *Transforming nursing education through problem-based learning* (pp. 51-74). Sudbury, MA: Jones and Bartlett Publishers.

Cyr, D., Childs, R., & Elgie, S. (2013). *Recruiting students for research in postsecondary education: A guide.* Toronto, ON: Higher Education Quality Council of Ontario.

Daniel, W.W., & Cross, C.L. (2013). *Biostatistics: A foundation for analysis in the health sciences* (10th ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Darling, N., Caldwell, L.L., & Smith, R. (2005). Participation in school-based extracurricular activities and adolescent adjustment. *Journal of Leisure Research, 37*(1), 51-*77.*

de Anda, D. de (2007). Intervention research and program evaluation in the school setting: Issues and alternative research designs. *Children & Schools, 29*(2), 87-94.

del Bueno, D. (1994). Why can't new grads think like nurses? *Nurse Educator, 19*(4), 9-11.

del Bueno, D. (2005). A crisis in critical thinking. *Nursing Education Perspectives, 26*(5), 278-282.

Des Jarlais, D.C., Lyles, C., Crepaz, N., & TREND Group. (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health, 94*(3), 361-366.

Dewey, J. (1938). *Experience & education.* New York, NY: Kappa Delta Pi.

DiCenso, A., Guyatt, G., & Ciliska, D. (2005). *Evidence-based nursing: A guide to clinical practice*. St. Louis, MO: Elsevier Mosby.

Dickert, N., & Grady, C. (1999). What's the price of a research subject? Approaches to payment for research participation. *The New England Journal of Medicine, 341*(3), 198-203.

Donahue, M.P. (1985). *Nursing: The finest art: An illustrated history*. St. Louis, MO: Mosby.

Dougherty, M.R.P., & Hunder, J.E. (2003). Probability judgment and subadditivity: The role of working emmory capacity and constraining retrieval. *Memory & Cognition, 31*, 968-982.

Dowie (1988). *Professional judgment and decision making: Introductory texts 1-4.* Milton Keynes, England: Open University Press.

Ebright, P.R., Patterson, E.S., Chalko, B.A., & Render, M.L. (2003). Understanding the complexity of registered nurse work in acute care settings. *Journal of Nursing Administration, 33*(12), 630-638.

Elliot, J., Rutty, C., & Villeneuve, M. (2013). *Canadian Nurses Association: One hundred years of service.* Ottawa, ON: Canadian Nurses Association.

Elstein, A.S. (1994). What goes around comes around: Return of the hypothetico-deductive strategy. *Teaching and Learning in Medicine, 6*(2), 121-123.

Elstein, A.S., & Bordage, G. (1988). Psychology of clinical reasoning. In J. Dowie & A. Elstein (Eds.), *Professional judgment: A reader in clinical decision making* (pp. 109-129). Cambridge, MA: Cambridge University Press.

Elstein, A.S., Shulman, L.S., & Sprafka, S.A. (1978). *Medical problem solving: An analysis of clinical reasoning.* Cambridge, MA: Harvard University Press.

Engel, G.L. (1977). The need for a new medical model: A challenge for biomedicine. *Science, 196*(4286), 129-136.

Ericsson, A.K., & Stasewski, J.J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 235-268). Hillesdale, NJ: Lawrence Erlbaum Associates, Inc.

Eriksson, K., & Lindström, U.A. (1997). Abduction – A way to deeper understanding of the world of caring. *Scandinavian Journal of Caring Sciences, 11*(4), 195-198.

Fann, K.T. (1970). *Peirce's theory of abduction.* The Hague, Netherlands: Martinus Nijhoff.

Fisher, S.D., Gettys, C.F., Manning, C., Mehle, T., & Baca, S. (1983). Consistency checking in hypothesis generation. *Organizational Behavior and Human Performance, 31*(2), 233-254.

Flin, R., Slaven, G., & Stewart, K. (1996). Emergency decision making in the offshore oil and gas industry. *Human Factors, 38*, 262-277.

Florin, J., Ehrenberg, A., & Ehnfors, M. (2005). Quality of nursing diagnoses: Evaluation of an educational intervention. *International Journal of Nursing Terminologies and Classifications, 16*(2), 33-43.

Fraser, S.W., & Greenhalgh, T. (2001). Coping with complexity: Educating for capability. *BMJ, 323* (October 6), 799-803.

Freire, P. (1970). *Pedagogy of the oppressed.* New York, NY: Herder and Herder.

Giuliano, K.K. (2003). Expanding the use of empiricism in nursing: Can we bridge the gap between knowledge and clinical practice? *Nursing Philosophy, 4*, 44-52.

Glaser, R., & Chi, M.T.H. (1988). Overview. In M.T.H. Chi, R. Glaser, & M.J. Farr (Eds.), *The nature of expertise* (pp. xv-xxviii). Hillsdale, NJ: Lawrence Erlbaum Associates.

Godfrey-Smith, P. (2003). *Theory and reality: An introduction to the philosophy of science.* Chicago, IL: The University of Chicago Press.

Gogtay, N.J. (2010). Principles of sample size calculation. *Indian Journal of Ophthalmology, 58*(6), 517-518.

Grady, C. (2001). Money for research participation: Does it jeopardize informed consent? *The American Journal of Bioethics, 1*(2), 40-44.

Guest, A., & Schneider, B. (2003). Adolescents' extracurricular participation in context: The mediating effects of schools, communities, and identity. *Sociology of Education, 76*, 89-105.

Haig, B.D. (1999). Construct validation and clinical assessment. *Behaviour Change, 16*(1), 64-73.

Haig, B.D. (2008). Scientific method, abduction, and clinical reasoning. *Journal of Clinical Psychology, 64*(9), 1013-1018.

Hamers, J.P.H., Abu-Saad, H.H., & Halfens, R.J.G. (1994). Diagnostic process and decision making in nursing: A literature review. *Journal of Professional Nursing, 10*(3), 154-163.

Harbison, J. (1991). Clinical decision making in nursing. *Journal of Advanced Nursing, 16*(4), 404-407.

Hartung, D.M. & Touchette, D. (2009). Overview of clinical research design. *American Journal of Health-System Pharmacy, 66*, 398-408.

Hay, D., & Kinchin, I. (2008). Using concept mapping to measure learning quality. *Education + Training, 50*(2), 167-182.

Hinck, S.S., & Brandell, M.E. (1999). Service learning: Facilitating academic learning and character development. *National Association of Secondary School Principals Bulletin, 83*, 16-25.

Hully, S.B., Cummings, S.R., Browner, W.S., Grady, D.G., & Newman, T.B. (2007). *Designing clinical research* (3rd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Ingram, C., Ray, K., Landeen, J., & Keane, D.R. (1998). Evaluation of an educational game for health sciences students. *Journal of Nursing Education, 37*(6), 240-246.

Johnson, T.R., Zhang, J., & Wang, H. (1997). A hybrid learning model of abductive reasoning. In R. Sun & F. Alexandre (Eds.), *Connectionist symbolic integration* (pp. 91-112). Hillsdale, NJ: Lawrence Erlbaum.

Kinchin, I.M., De-Leij, F.A.A.M., & Hay, D.B. (2005). The evolution of a collaborative concept mapping activity for undergraduate microbiology students. *Journal of Further and Higher Education, 29*(1), 1-14.

Kinnear, J.L. (1994). The professionalization of Canadian nursing, 1924-32: Views in the CN and the CMAJ. *Canadian Bulletin of Medical History, 11*, 153-174.

Klein, G.A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In J. Orasanu (Ed.), *Decision making in action: Models and methods* (pp. 138-147). Westport, CT: Ablex Publishing.

Knowles, M. (1975). *Self-directed learning: A guide for students and teachers.* New Jersey: Prentice Hall.

Knox, R., Seth, R., McElveen, N., Bergstein, M., & Longo, J. (2003). *Short term memory based on gender.* Retrieved from http://andrewd.ces.clemson.edu/courses/cpsc412/fall03/teams/reports/group5.pdf

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.

Koutoukidis, G., Stainton, K., & Hughson, J. (2013). *Tabbner's nursing care: Theory and practice* (6th ed.). Chatswood, NSW: Elsevier.

Kwon, Y.J., Jeong, J.S., & Park, Y.B. (2006). Roles of abductive reasoning and prior belief in children's generation of hypotheses about pendulum motion. *Science & Education, 15*, 643-656.

Kwon, Y.J., Lee, J.K., Shin, D.H., & Jeong, J.S. (2009). Changes in brain activation induced by the training of hypothesis generation skills: An fMRI study. *Brain and Cognition, 69*(2), 391-397.

Landers, R.N. (2013). *A step-by-step introduction to statistics for business.* London, UK: SAGE Publications Ltd.

Lawson, A.E., & Daniel, E.S. (2011). Inferences of clinical diagnostic reasoning and diagnostic error. *Journal of Biomedical Informatics, 44*(3), 402-412.

Lenth, R.V. (2007). Post hoc power: Tables and commentary. Technical report No. 378. Department of Statistical and Actuarial Science, University of Iowa. Retrieved January 28, 2015, from http://www.stat.uiowa.edu/files/stat/techrep/tr378.pdf.

Levett-Jones, T., Hoffman, K., Dempsey, J., Jeong, S.Y., Noble, D., Norton, C.A., ... Hickey, N. (2010). The 'five rights' of clinical reasoning: An educational model to enhance nursing students' ability to identify and manage clinically 'at risk' patients. *Nursing Education Today, 30*(6), 515-520.

Levin, R.F., Lunney, M., & Krainovich-Miller, B. (2004). Improving diagnostic accuracy using an evidence-based nursing model. *International Journal of Nursing Terminologies and Classifications, 15*(4), 114-122.

Li, S., & Kenward, K. (2006). A national survey of nursing education and practice of newly licensed nurses. *JONA's Healthcare Law, Ethics, and Regulation, 8*(4), 110-115.

Lipscomb, M. (2012). Abductive reasoning and qualitative research. *Nursing Philosophy, 13*, 244-256.

Loiselle, C. G., Profetto-McGrath, J., Polit, D. F., & Beck, C. T. (2011). *Canadian essentials of nursing research* (3rd ed.). New York, NY: Wolters Kluwer Health | Lippincott, Williams & Wilkins.

Lunny, M. (2001). *Critical thinking and nursing diagnosis: Case studies and analyses.* Philadelphia, PA: NANDA (North American Nursing Diagnosis Association).

Lunny, M., (2008). Critical need to address accuracy of nurses' diagnoses. *Online Journal of Issues in Nursing, 13*(1).

Mann, J. (2012). Critical thinking and clinical judgment skill development in baccalaureate nursing students. *The Kansas Nurse, 87*(1), 26-31.

Margetson, D. (1997). Why is problem-based learning a challenge? In D. Boud & G. Feletti (Eds.), *The challenge of problem-based learning* (2nd ed., pp. 36-44). London, UK: Kogan Page Ltd.

Markert, R.J., O'Neill, S.C., & Bhatia, S.C. (2003). Using a quasi-experimental research design to assess knowledge in continuing medical education programs. *The Journal of Continuing Education in the Health Professions, 23*, 157-161.

Marriner-Tomey, A., & Alligood, M.R. (2006). *Nursing theorists and their work* (6th ed.). St. Louis, MO: Mosby-Elsevier.

Marsh, H.W., & Kleitman, S. (2002). Extracurricular activities: The good, the bad, and the nonlinear. *Harvard Educational Review, 72*(4), 464-512.

Mayer, R.E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59*(1), 14-19.

McCarthy, M. (1981). Nursing process – application of current thinking in problem solving. *Journal of Advanced Nursing, 6*(3), 173-177.

McClelland, J., Rumelhar, D., & the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructures of cognition.* Cambridge, MA: MIT Press.

McGonigle, D., & Mastrian, K.G. (2012). *Nursing informatics and the foundation of knowledge* (2nd ed.). Burlington, MA: Jones & Bartlett Learning.

McMaster University. (2015). *Admission requirements.* Retrieved from http://future.mcmaster.ca/admission/admission-requirements

McMaster University. (2013a). *N1F03: Introduction to nursing and health I: Tutor Guide.* Hamilton, ON: McMaster Mohawk Conestoga BScN Program.

McMaster University. (2013b). *N1G03: Introduction to nursing and health II: Tutor Guide.* Hamilton, ON: McMaster Mohawk Conestoga BScN Program.

McMaster University. (2013c). *N2MM3: Nursing concepts in health and illness I: Tutor Guide.* Hamilton, ON: McMaster Mohawk Conestoga BScN Program.

McMaster University. (2013d). *N2NN3: Nursing concepts in health and illness II: Tutor Guide.* Hamilton, ON: McMaster Mohawk Conestoga BScN Program.

McMaster University BScN Handbook. (2012). *Undergraduate nursing education: Program Handbook.* Hamilton, ON: McMaster University.

Miller, G.E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine, 65*(9), S63-S67.

Mohawk College (2015). *Nursing (BScN) admission requirements: Minimum requirements for admission to the nursing (BScN) degree program.* Retrieved from http://www.mohawkcollege.ca/health-sciences-programs/nursing-bscn-degree/admission-requirements.html

Mirza, N., Akhtar-Danesh, N., Noesgaard, C., Martin, L., & Staples, E. (2014a). A concept analysis of abductive reasoning. *Journal of Advanced Nursing, 70*(9), 1980-1994.

Mirza, N., Akhtar-Danesh, N., Staples, E. Martin, L., & Noesgaard, C. (2014b). Comparative analysis of external validity reporting in non-randomized intervention studies. *Canadian Journal of Nursing Research, 46*(4), 47-64.

Monti, E.J., & Tingen, M.S. (1999). Multiple paradigms of nursing science. *Advances in Nursing Science, 21*(4), 64-80.

NANDA (North American Nursing Diagnosis Association). (2014). *1980-1989.* Retrieved from http://www.nanda.org/nanda-international-history-1980-89.html

Needleman, J. & Buerhaus, P. (2003). Nurse staffing and patient safety: Current knowledge and implications for action. *International Journal of Quality in Health Care, 15*(4), 275-277.

Newell, A., & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Nightingale, F. (1859 reprinted 1946). *Notes on nursing: What it is and what it is not*. Philadelphia, PA: J.B. Lippincott.

Norman, G.R., Brooks, L.R., Colle, C.L., & Hatala, R.M. (1999). The benefit of diagnostic hypotheses in clinical reasoning: Experimental study of an instructional intervention for forward and backward reasoning. *Cognition and Instruction, 17*(4), 433-448.

Norman, G.R., Trott, A.D., Brooks, L.R., & Smith, E.K.M. (1994). Cognitive differences in clinical reasoning related to postgraduate training. *Teaching and Learning in Medicine, 6*(2), 114-120.

Oke, J.K. (2009). *Management information systems* (9th ed.). Pune, India: Nirali Prakashan.

Patel, V.L., Arocha, J.F., & Zhang, J. (2005). Thinking and reasoning in medicine. In: K.J. Holyoak & R.G. Morrison. *The Cambridge handbook of thinking and reasoning* (pp.727-750). New York, NY: Cambridge University Press.

Patel, V.L., & Groen, G.J. (1986). Knowledge-based solution strategies in medical reasoning. *Cognitive Science, 10*, 91-116.

Paul, R., & Elder, L. (2014). *The miniature guide to critical thinking concepts and tools* (7th ed.). Tomales, CA: Foundation for Critical Thinking.

Pavlov, I. (1927). *Conditioned reflexes.* London, UK: Oxford University Press.

Peirce, C.S. (1903 reprinted 1998). Pragmatism as the logic of abduction. In the Peirce Edition Project (Ed.), *The essential Peirce: Selected philosophical writings: Volume 2 (1893-1913)* (pp. 226-241). Bloomington, IN: Indiana University Press.

Peirce C.S. (1931-1958). *Collected papers of Charles Sanders Peirce*. Vol. 1-6, C. Hartshorne & P. Weiss (eds.). Vol. 7-8, A.W. Burks (ed.). Cambridge, MA: Harvard University Press.

Perkins, D.N. (1987). Thinking frames: An integrative perspective on teaching cognitive skills. In J. B. Baron & R.J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 41-61). New York, NY: Freeman & Company.

Pettigrew, A. C. (2015). Learning and students. In M.H. Oermann (Ed.), *Teaching in nursing and role of the educator: The complete guide to best practice in teaching, evaluation and curriculum development* (pp. 15-33). New York, NY: Springer.

Polit, D.F., & Beck, C.T. (2012). *Nursing research: Generating and assessing evidence for nursing practice* (9th ed.). Philadelphia, PA: Wolters Kluwer Health | Lippincott Williams & Wilkins.

Popper, K. (1945 reprinted 2013). *The open society and its enemies.* Princeton, NJ: Princeton University Press.

Purling, A., & King, L. (2012). A literature review: Graduate nurses' preparedness for recognising and responding to the deteriorating patient. *Journal of Clinical Nursing, 21*, 3451-3465.

Raholm, M.B. (2010a). Abductive reasoning and the formation of scientific knowledge within nursing research. *Nursing Philosophy, 11*, 260-270.

Raholm, M.B. (2010b). Theory development and the logic of discovery. *International Journal of Human Caring, 14*(3), 36-42.

Reed, G.P. (1995). A treatise on nursing knowledge development for the 21st century: Beyond postmodernism. *Advances in Nursing Science, 17*(3), 70-84.

Rideout, E. (2001a). Evaluating student learning. In E. Rideout (Ed.), *Transforming nursing education through problem-based learning* (pp. 215-237). Sudbury, MA: Jones and Bartlett Publishers.

Rideout, E., (2001b). *Transforming nursing education through problem-based learning*. Sudbury, MA: Jones and Bartlett Publishers.

Rideout, E., & Carpio, B. (2001). The problem-based learning model of nursing education. In E. Rideout (Ed.), *Transforming nursing education through problem-based learning* (pp. 21-49). Sudbury, MA: Jones and Bartlett Publishers.

Roediger III, H.L., & Karpicke, J.D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*(3), 181-210.

Rolfe, G. (1997). Science, abduction and the fuzzy nurse: An exploration of expertise. *Journal of Advanced Nursing, 25*(5), 1070-1075.

Romyn, D.M. (2001). Disavowal of the behaviorist paradigm in nursing education: What makes it so difficult to unseat? *Advances in Nursing Science, 23*(3), 1-10.

Roth, D. (1996). A connectionist framework for reasoning: Reasoning with examples. In *Proceedings of the thirteenth national conference on artificial intelligence – Volume 2* (pp. 1256-1261). Portland, OR: AAAI Press.

Round, A.P. (1999). Teaching clinical reasoning – a preliminary controlled study. *Medical Education, 33*, 480-483.

Ruiz-Primo, M. (2000). On the use of concept maps as an assessment tool in science: What we have learned so far. *Revista Electronica de Investigacion Educativa, 2*(1). Retrieved April 10, 2014 from: http://redie.uabc.mx/vol2no1/contents-ruizpri.html

Scheckel, M. (2009). Nursing education: Past, present, future. In G. Roux & J.A. Halstead (Eds.), *Issues and trends in nursing: Essential knowledge for today and tomorrow* (pp.27-61). Sudbury, MA: Jones and Bartlett Publishers.

Schoenberg, N.E., Bardach, S.H., Manchikanti, K.N., & Goodenow, A.C. (2011). Appalachian residents' experiences with and management of multiple morbidity. *Qualitative Health Research, 21*(5), 601-611.

Schunk, D.H. (2012). *Learning theories: An educational perspective*. Boston, MA: Pearson.

Silén, C. (2004). Does problem-based learning make students "go meta"? In M. Savin-Baden & K. Wilkie (Eds.), *Challenging research in problem-based learning* (pp. 144-155). London, UK: Open University Press.

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

Simmons, B., Lanuza, D., Fonteyn, M., Hicks, F., & Holm, K. (2003). Clinical reasoning in experienced nurses. *Western Journal of Nursing Research, 25*(6), 701-719.

Singer, E., & Bossarte, R.M. (2006). Incentives for survey participation: When are they 'coercive'? *American Journal of Preventative Medicine, 31*(5), 411-418.

Skinner, B.F. (1953). *Science and human behavior.* New York, NY: Free Press.

Sumathi, S., & Saravanavel, P. (2008). *Marketing research & consumer behaviour.* New Delhi, India: Vikas Publishing House.

Tanner, C.A. (1982). Instruction on the diagnostic process: An experimental study. In M.J. Kim & D.A. Mortiz (Eds.), *Classification of nursing diagnosis: Proceedings of the 3rd and 4th National Conference* (pp. 145-152). New York, NY: McGraw-Hill.

Tanner, C.A. (2006). Thinking like a nurse: A research-based model of clinical judgment in nursing. *Jounral of Nursing Education, 45*(6), 204-211.

Tanner, C.A. (2007). The curriculum revolution revisited. *Journal of Nursing Education,46*(2), 51-52.

Tanner, C.A., Padrick, K.P., Westfall, U.E., & Putzier, D.J. (1987). Diagnostic reasoning strategies of nurses and nursing students. *Nursing research, 36*(6), 358-363.

The Committee on Education of the National League of Nursing Education. (1918). *Standard curriculum for schools of nursing*. Baltimore, MD: The Waverly Press.

Thomas, R.P., Dougherty, M.R., Sprenger, A.M., & Harbison, J.I. (2008). Diagnostic hypothesis generation and human judgment. *Psychological Review, 115*(1), 155-185.

Thompson, C. (1999). A conceptual treadmill: The need for 'middle ground' in clinical decision making theory in nursing. *Journal of Advanced Nursing, 30*(5), 1222-1229.

Thompson, C., & Dowding, D. (2002). Decision-making and judgement in nursing – an introduction. In C. Thompson & D. Dowding (Eds.), *Clinical decision-making and judgement in nursing* (pp. 1-19). Edinburgh: Churchill Livingstone.

Thorndike, E.L. (1911). *Animal intelligence.* New York, NY: Macmillan.

Tomic, W. (1993). Behaviorism and cognitivism in education. *Psychology, A Journal of Human Behavior, 30*(3/4), 38-46.

Tyler, R. (1949). *Basic principles of curriculum and instruction.* Chicago, IL: University of Chicago Press.

Vandewater, D.A. (2004). *Best practices in competence assessment of health professionals.* Halifax, NS: College of Registered Nurses of Nova Scotia.

van der Vleuten, C.P.M. (2000). Validity of final examinations in undergraduate medical training. *British Medical Journal, 321*, 1217-1219.

Vanderbilt University Department of Biostatistics. (2014). *PS: Power and sample size calculation.* Retrieved from: http://biostat.mc.vanderbilt.edu/PowerSampleSize

Vertue, F.M., & Haig, B.D. (2008). An abductive perspective on clinical reasoning and case formulation. *Journal of Clinical Psychology, 49*(9), 1046-1068.

Vess, P.J. (2007). *Competence assessment methods in advanced practice nursing: A synthesis of the literature* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses A&I. (UMI Microform 3296679).

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes.* (M. Cole, V. John-Steiner, S. Scribner, and E. Souberman, Eds.). Cambridge, MA: Harvard University Press.

Wang, H., Johnson, T.R., & Zhang, J. (2006). A hybrid system of abductive tactical decision making. *International Journal of Hybrid Intelligence Systems, 3*, 23-33.

Ward, T., & Haig, B. (1997). Abductive reasoning and clinical assessment. *Australian Psychologist, 32*(2), 93-100.

Wass, V., Van der Vleuten, C., Shatzer, J., & Jones, R. (2001). Assessment of clinical competence. *The Lancet, 357*(9260), 945-949.

Weiser, M., & Shertz, J. (1983). Programming problem representation is novice and expert programmers. *Instructional Journal of Man-Machine Studies, 14,* 391-396.

Wilkie, K. (2004). Becoming facilitative: Shifts in lecturers' approaches to facilitating problem-based learning. In M. Savin-Baden & K. Wilkie (Eds.), *Challenging research in problem-based learning* (pp. 81-92). London, UK: Open University Press.

Winter, J.C.F., & Dodou, D. (2011). Predicting academic performance in engineering using high school exam scores. *International Journal of Engineering Education, 27*(6), 1343-1351.

Wolf, F.M., Gruppen, L.D., & Billie, J.E. (1988). Use of the competing-hypotheses heuristic to reduce 'pseudodiagnosticity'. *Journal of Medical Education, 63*(7), 548-554.

Wolff, A.C., Regan, S., Pesut, B., & Black, J. (2010). Ready for what? An exploration of the meaning of new graduate nurses' readiness for practice. *International Journal of Nursing Education Scholarship, 7*(1).

Wong, T.K.S. & Chung, J.W.Y. (2002). Diagnostic reasoning processes using patient simulation in different learning environments. *Journal of Clinical Nursing, 11*, 65-72.

Woods, N.N., Brooks, L.R., & Norman, G.R. (2007). The role of biomedical knowledge in diagnosis of difficult clinical cases. *Advances in Health Sciences Education, 12*(4), 417-426.

Appendix 1: A Concept Analysis of Abductive Reasoning

**JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS**

Mar 05, 2015

This Agreement between Noeman A Mirza ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	3582410844682
License date	Mar 05, 2015
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Journal of Advanced Nursing
Licensed Content Title	A concept analysis of abductive reasoning
Licensed Content Author	Noeman A. Mirza,Noori Akhtar-Danesh,Charlotte Noesgaard,Lynn Martin,Eric Staples
Licensed Content Date	Mar 12, 2014
Pages	15
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Title of your thesis / dissertation	Effects of Abductive Reasoning Training on Hypothesis Generation Abilities of First and Second Year Baccalaureate Nursing Students
Expected completion date	Jun 2015
Expected size (number of pages)	200
Requestor Location	Noeman A Mirza 1275 Danforth Ave. Suite 412 Toronto, ON M4J5B6 Canada Attn: Noeman A Mirza
Billing Type	Invoice
Billing Address	Noeman A Mirza 1275 Danforth Ave. Suite 412 Toronto, ON M4J5B6 Canada Attn: Noeman A Mirza
Total	0.00 CAD
Terms and Conditions	

Abstract

Aim

To describe an analysis of the concept of abductive reasoning.

Background

In the discipline of nursing, abductive reasoning has received only philosophical attention and remains a vague concept. In addition to deductive and inductive reasoning, abductive reasoning is not recognized even in prominent nursing knowledge development literature. Therefore, what abductive reasoning is and how it can inform nursing practice and education was explored.

Design

Concept analysis.

Data Sources

Combinations of specific keywords were searched in Web of Science, CINAHL, PsychINFO, PubMed, Medline and EMBASE. The analysis was conducted in June 2012 and only literature before this period was included. No time limits were set.

Methods

Rodger's evolutionary method for conducting concept analysis was used.

Results

Twelve records were included in the analysis. The most common surrogate term was retroduction while related terms included intuition and pattern and similarity recognition. Antecedents consisted of a complex, puzzling situation and a clinician with creativity, experience and knowledge. Consequences included the formation of broad

hypotheses which enhance understanding of care situations. Overall, abductive reasoning was described as the process of hypothesis or theory generation and evaluation. It was also viewed as inference to the best explanation.

Conclusion

As a new approach, abductive reasoning could enhance reasoning abilities of novice clinicians. It can not only incorporate various ways of knowing, its holistic approach to learning also appears to be promising in problem-based learning. Since nursing literature on abductive reasoning is predominantly philosophical, practical consequences of abductive reasoning warrant further research.

Key Words

Abductive reasoning, hypothesis generation, hypothetico-deductive method, clinical reasoning, abduction, problem-based learning, ways of knowing, concept analysis.

Summary Statement

Why is this research or review needed?

- Prominent nursing literature on theory and knowledge development discusses only deductive and inductive reasoning while discussion on abductive reasoning is lacking.
- Abductive reasoning has received limited philosophical attention in nursing and the concept and its practical implications remain vague.

What are the key findings?

- Abductive reasoning is the process of generating hypotheses, theories or explanations and precedes deductive and inductive reasoning.
- To engage in abductive reasoning, there must be a complex or puzzling situation and a clinician who possesses creativity, experience and knowledge.
- Abductive reasoning results in the formation of broad hypotheses which lead to an enhanced understanding of the care situation.

How should the findings be used to influence policy/practice/research/education?

- Abductive reasoning can be introduced as a new approach in problem-based learning which presently uses hypothetico-deductive reasoning.
- Abductive reasoning is conducive to different ways of knowing and can allow practitioners to gain a broader and deeper understanding of a care situation.
- Practicing nurses can utilize abductive reasoning to explain nursing-related issues for newly admitted clients which could guide further care planning.

Introduction

Abductive reasoning is a form of synthetic inference through which meaningful underlying patterns of selected phenomena are recognized to comprehend a complex reality and expand scientific knowledge (Raholm 2010a, Raholm 2010b). According to Eriksson and Lindström (1997), abductive reasoning guides the generation of hypotheses, the consequences of which are explicated logically through deductive reasoning and empirically through inductive reasoning. Reasoning associated with clinical practice, or clinical reasoning, has been declared to be utilizing a hypothetico-deductive method which promotes deductive and inductive approaches to reasoning (Simmons 2010), while an abductive approach is lacking.

After recognizing the importance of abductive reasoning as a first stage of inquiry and noticing its absence in the current hypothetico-deductive method of clinical reasoning, nursing scholars have made attempts to introduce the idea of abductive reasoning in nursing literature as a new way of thinking about clinical practice and clinical reasoning (Raholm 2010a, Eriksson & Lindström 1997, Reed 1995, Rolfe 1997). However, their accounts of abductive reasoning either vary from one another, lack depth in elucidating the entire process of abductive reasoning, or present very brief interdisciplinary comparisons often incorporating philosophical concepts into nursing or comparing nursing to medicine. Since these limited discussions on abductive reasoning and its relation to clinical nursing practice fall short in capturing the abductive reasoning process in its entirety, a comprehensive interdisciplinary exploration of abductive reasoning as a concept is warranted.

Reported in this paper, is a concept analysis of abductive reasoning. The purpose of this analysis is to determine what abductive reasoning is, how it differs from other forms of reasoning such as deduction and induction in relation to clinical reasoning and how it is understood in the context of clinical practice in the disciplines of nursing, medicine and psychology. To the best of our knowledge, a concept analysis of abductive reasoning has not been performed before in the literature. Therefore, this analysis aims to enrich understanding of abductive reasoning while offering direction for further research which can contribute to nursing practice and education. Since abductive reasoning is a developing concept which has been drawing interdisciplinary debate, Rodger's evolutionary and inductive method of concept analysis is used (Rodgers 2000, Tofthagen & Fagerstrøm 2010).

While previous approaches to concept analysis value reduction in an attempt to isolate the essence of a concept from its dynamic interrelationships with the world (Chinn & Jacobs 1983, Smith & Medin 1981, Walker & Avant 1983), Rodgers (1989)'s method to analysis is founded on the assumption that concepts are continually changing and do not remain constant across contexts. This makes it an ideal choice for exploring abductive reasoning, which is still evolving as a concept in several disciplines including music, art, mathematics and information technology. For the purpose of systematic reporting, the concept analysis on abductive reasoning is divided into four major sections: 1) background, 2) data collection, 3) results and analysis and 4) discussion. Each of these sections is further divided into sub-sections that correspond with the evolutionary phases outlined by Rodgers' concept analysis method.

Background

Emergence of Clinical Reasoning

Rimoldi (1961) perhaps conducted the first experiment that compared diagnostic reasoning abilities of students and clinicians. He found that expert physicians had increased ability in selecting relevant data and narrowing diagnostic possibilities. This was followed by Newell and Simon (1972)'s work on problem-solving, after which research focusing on information-processing psychology began to flourish. A few years later, Elstein *et al.* (1978) investigated clinical competency. This influenced the development of the *hypothetico-deductive* method of reasoning in medicine, which became the established approach of diagnostic reasoning that allowed clinicians to think like scientists (Coderre *et al.* 2003).

The hypothetico-deductive method also made its way into clinical psychology and nursing practice. In their work, Ward and Haig (1997) declare that models of clinical reasoning found in the psychology and behavioural literature employ the hypothetico-deductive method to clinical reasoning. In nursing, not only do nurses use the hypothetico-deductive method when making clinical decisions (Gordon 1987, McFadden & Gunnett 1992, Radwin 1989, Tanner 1982, Tanner *et al.* 1987, White *et al.* 1992), their education immerses them in a form of clinical reasoning that is rooted in the hypothetico-deductive method (Wong & Chung 2002). In her concept analysis of clinical reasoning, Simmons (2010) claims that clinical reasoning in nursing is founded on the hypothetico-deductive method. She further mentions that clinical reasoning in nursing relies on both deductive and inductive reasoning. However, Simmons does not discuss abductive reasoning.

Hypothetico-Deductive Reasoning

While ‘deductive reasoning’ (i.e., deduction) derives a particular conclusion from a general premise, ‘inductive reasoning’ (i.e., induction) derives a general conclusion from a set of particular statements (Patel *et al.* 2005). The hypothetico-deductive method focuses on hypothesis-testing and consists of deduction which is initiated by a hypothesis (hence the name ‘hypothetical’ or ‘hypothetico’ deduction). For example, when faced with a particular clinical situation, a clinician following the hypothetico-deductive method develops a set of hypotheses based on previous knowledge or recognition of patterns. She then tests these hypotheses indirectly by collecting data which either confirm or disconfirm the predictions in her hypotheses. The clinician then reasons inductively when she extracts conclusions from the results of her hypothesis-testing efforts to consider and implement various treatment or intervention options to deal with the presented clinical situation or client complaint (Ward & Haig 1997).

The methods employed by clinicians as part of clinical reasoning either emphasize the process of induction or hypothetico-deduction for explaining clinical situations (Patel *et al.* 2005, Simmons 2010). However, both forms of reasoning fail to describe the initial phase of inquiry which is related to the discovery of hypotheses (Raholm 2010a, Haig 1999, Ward & Haig 1997). This form of inference is referred to as ‘abductive reasoning’ (i.e., abduction), which involves both the generation and refinement of explanatory hypotheses which is a pre-requisite for deductive reasoning (Eriksson & Lindström 1997).

Abductive Reasoning

Abduction, deduction and induction were originally derived from the work of Aristotle and reintroduced in present times by the American philosopher and father of pragmatism, Charles Sanders Peirce (Eriksson & Lindström 1997, Raholm 2010a). Dissatisfied with the explanation that hypotheses are mere guesses and do not have a logic of discovery, Peirce’s purpose was to uncover the logic through which new ideas come into existence (Fann 1975). He called this stage abduction and described it as the first stage of inquiry where new ideas or hypotheses are invented to explain meaningful underlying patterns of selected phenomena (Peirce 1903; Peirce 1931-1958). Since hypotheses are still plausible at this first stage of inquiry, Peirce (1931-1958) recommends they be further explicated logically through deductive reasoning and empirically through inductive reasoning.

Abductive reasoning is a creative inference which involves integration and justification of ideas to develop new knowledge. While abductive reasoning allows one to conceive ideas from vague, possible, or potentially possible phenomena, deductive and inductive reasoning allow for the consequent processing of those ideas (Raholm 2010a). This process can be explained as such: 1) Surprising phenomena emerge and require an explanation because they do not follow an accepted hypothesis; 2) a new hypothesis that predicts these phenomena is adopted through abduction; 3) necessary and probable experimental consequences of the hypothesis are traced out through deduction; 4) when tests verify prediction after prediction, the hypothesis is stationed among scientific results through induction (Haig 1999, Raholm 2010a, Lawson & Daniel 2011).

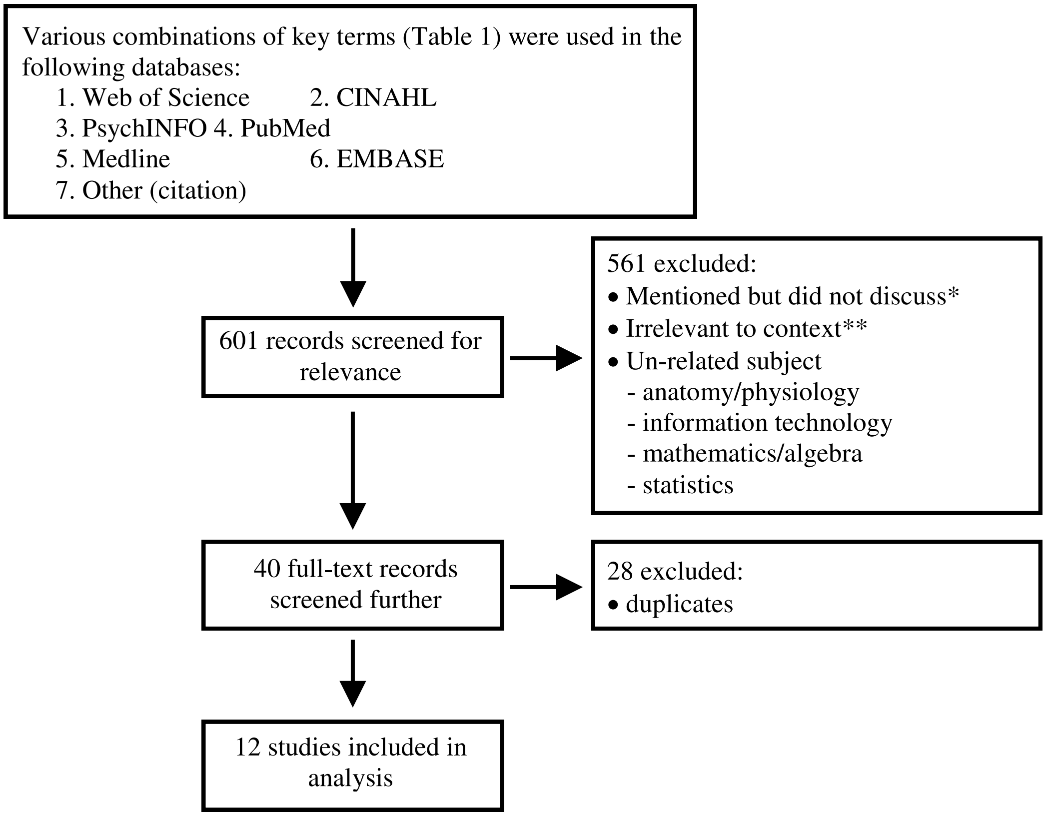
Data Sources

Several databases and search strategies were employed to review the literature and yield an ample number of records for analysis. The concept of ‘abductive reasoning’ and its synonym, abduction, were constantly used as major keywords. Since the search was related to ‘nursing’, ‘health’ and ‘caring’ in the context of ‘clinical reasoning’ and ‘clinical practice’, these terms were also included in the search strategy as keywords. Combinations of these terms were searched in several databases that focused on literature from clinical disciplines of nursing, medicine and psychology. These included Web of Science, CINAHL, PsychINFO, PubMed, Medline and EMBASE. The search strategy is presented in Figure A1.1 and Table A1.1. The analysis was conducted in June 2012 and time limits were not set since a previous preliminary search using a combination of the abovementioned keywords generated only a handful of records.

In addition to the term ‘abductive reasoning’, ‘abduct*’ was also used to indicate abduct, abductive and abduction. These terms were often combined with the term ‘clinical’ and/or ‘reasoning’ to avoid results focusing on the anatomical definition of abduction and to facilitate the retrieval of sources focusing on the cognitive thinking process. Due to the low number of results in nursing literature, ‘artificial intelligence’ was also used in combination with ‘abduction’ in CINAHL to maximize the opportunity for retrieving more nursing-based articles on abductive reasoning. However, this attempt did not yield any results. In instances where the use of clinical failed to generate sufficient results, other terms such as ‘caring’ or ‘health’ were utilized.

Table A1.1: Search Strategy

Database	**Records Found**	**Irrelevant Records**	**Relevant Records**
Web of Science	17	10	7
1. 'abductive reasoning' and 'clinical'			
2. 'abduction' and 'clinical reasoning'			
3. 'abduction' and 'clinical' and 'reasoning'			
CINAHL	53	44	9
1. 'abductive reasoning'			
2. 'abduction' and 'artificial intelligence'			
3. 'abductive' in Title			
4. 'abduct*' and 'reason' in Abstracts			
PsychINFO	54	45	9
1. 'abductive reasoning' and 'clinical'			
2. 'abductive reasoning' in Abstracts			
3. 'abduction' and 'clinical reasoning'			
PubMed	188	184	4
1. 'abductive reasoning' and 'clinical'			
2. 'abduction' and 'clinical reasoning'			
3. 'abduction and clinical' and 'nursing or caring or health'			
Medline	95	91	4
1. 'abductive reasoning'			
2. 'abduction' and 'clinical reasoning'			
3. 'abduction and clinical' and 'nursing or caring or health'			
EMBASE	193	187	6
1. 'abductive reasoning' and 'clinical'			
2. 'abduction' and 'clinical reasoning'			
3. 'abduction and clinical' and 'nursing or caring or health'			
Other Methods			
1. Located through citation	1	0	1
Total records	**601**	**561**	**40**
Duplicate records removed			**28**
Records included in analysis			**12**



*Abduction/abductive reasoning briefly mentioned or referenced but not thoroughly discussed as a process of reasoning as related to decision-making, problem-solving and clinical judgment

**Irrelevant to context of clinical reasoning (i.e., abduction/abductive reasoning discussed in general context rather than a clinical reasoning or clinical practice context)

Figure A1.1. Flow diagram of search strategy.

A total of 601 records were found with the search terms, while one of these was found through citation. Abstracts of these records were read to determine their relevance to the concept analysis. On occasion, full-text of some articles was also explored. Of the 601 records, 561 were excluded as they were irrelevant to the purpose of the concept analysis. Many of these records were excluded because they did not discuss abductive reasoning or abduction in relation to clinical reasoning, clinical practice, or any activity associated with the clinical context; while some were excluded because of their focus on disciplines other than the three identified. From the remaining 40 records, 28 were found to be duplicates and were also excluded. This resulted in a total number of 12 records which were included in the concept analysis.

Results

The twelve articles included in the analysis were read once to obtain a general understanding of how abductive reasoning was understood in the specified context. All articles were then re-read and critically analyzed. Data were extracted from these articles using a tool which was based on the various items of Rodgers (1989)'s evolutionary method of concept analysis. These allowed for the identification of the concept's surrogate and related terms, antecedents, consequences, defining attributes, examples and a model case. According to Rodgers, extraction of data based on these items helps provide key information of what the chosen concept is and is not.

According to Rodgers (1989), 'surrogate terms' have the same meaning as the identified concept while 'related terms' have different characteristics but maintain some common similarity with the concept. Rodgers defines 'antecedents' as events or phenomena that precede an instance of the chosen concept while 'consequences' are events or phenomena that follow an occurrence of the concept. She further describes 'defining attributes' as characteristics of the concept and 'examples' as the concrete instances described in the data sources that discuss the concept. The development of a 'model case' which is an everyday example that validates the concept and its characteristics is the end-result of the previously outlined steps.

Table A1.2: Application of Rogers' Method of Concept Analysis to Abductive Reasoning

Article	Context	Surrogate Terms	Related Terms	Antecedents	Attributes	Consequences	Examples
Psychology							
Haig (1999)	Construct validation and clinical assessment	Abduction, retroduction, abductive inference	Exploratory factor analysis, explanatory coherence	Surprising phenomena	Hypothesis/theory generation, inference to the best explanation	Explanation of evidence	Jury and crime detection generates theories
Haig (2008)	Clinical reasoning	Abduction	Exploratory factor analysis, analogical modeling, explanatory coherence	Cognitive ability	Hypothesis generation, inference to the best explanation	Plausible explanation, hypothesis	*
Vertue & Haig (2008)	Clinical reasoning and case formulation	Abductive theory of method	Exploratory factor analysis, data-driven reasoning, forward reasoning, bottom-up reasoning	Data collection	Phenomena detection, inferring causal mechanisms, developing and evaluating the causal model, formulating the case, explanatory coherence	Narrative, case, breadth	*
Ward & Haig (1997)	Clinical assessment	Abductive method/model	*	Phenomena detection	Explanatory theory, hypothesis/theory generation	Case	*

Article	Context	Surrogate Terms	Related Terms	Antecedents	Attributes	Consequences	Examples
Nursing							
Eriksson & Lindström (1997)	Nursing knowledge development	Abduction, abductive method, retroduction, abductive thinking, abductive inference, retroductive inference	*	Complex patterns of reality, broad and deep theoretical base, strong knowledge base	Hypothesis generation, recognition, interpretation, pattern seeking/ creating/renewing	Explanatory hypothesis, explanation	*
Raholm (2010a)	Scientific knowledge formation	Abduction, abductive model, retroduction, abductive logic, abductive inference	*	Previous knowledge, deep theoretical basis, creativity, boldness, tacit knowledge	Instinct, hypothesis invention, creative insight	Hypothesis, deduction, induction	Detectives anticipate disturbing phenomena yet seek explanation
Raholm (2010b)	Theory development	Abduction, synthetic reasoning	*	Surprising fact, puzzling data	Hypothesis creation,	Deeper knowledge and understanding, scientific progress, deduction	*
Reed (1995)	Knowledge development	Abduction, retroductive reasoning	*	Phenomena, experience, beliefs, pre-existing conceptual and empirical knowledge of patterns	Generation of theories and educated guesses	Deduction, induction, empirical testing	*
Rolfe (1997)	Clinical expertise	Abduction, fuzzy logic	Intuition, similarity recognition	Tacit knowledge	Information utilization, fuzzy (vague, imprecise) reasoning	Explanation	Computer self-drives helicopter

Article	Context	Surrogate Terms	Related Terms	Antecedents	Attributes	Consequences	Examples
Medicine							
Lawson & Daniel (2011)	Clinical diagnostic reasoning and diagnostic error	Abduction, If/then/therefore reasoning, retroduction	Analogical modeling/ reasoning/ transfer/ inference	Declarative knowledge, puzzling symptoms	Hypothesis generation	Explanation	Patient continues to have symptoms of gallbladder stones even after surgery
Magnani (1997)	Clinical reasoning and medical education	Abduction	*	Established medical knowledge, previous similar experience	Visual abduction, creative abduction, selective abduction, inference to the best explanation, uncertainty	Diagnostic hypothesis, decision-making	*
Upshur (1997)	Clinical reasoning and evidence-based medicine	Abduction	*	Surprising data	Inference to the best explanation	Explanation	*

*Not described in the article

A total of twelve articles ranging from 1997 – 2011 were reviewed for the concept analysis. Of these, five were from the discipline or nursing, while three were from medicine and four from psychology. Extracted data are presented in Table A1.2 based on specific items outlined by Rodgers (1989)'s evolutionary method of concept analysis. These items are also discussed individually in the subsequent sections and relevant interdisciplinary comparisons are made. After examining all twelve articles, saturation within and between disciplines was achieved regarding the characteristics of abductive reasoning as a concept within the context of clinical reasoning and clinical practice. However, some outliers were also identified, which are described and analyzed further with literature support.

Surrogate Terms

Several terms were interchangeably used with abductive reasoning. Of these, most had the same root word such as 'abduction', 'abductive method', 'abductive model', 'abductive thinking', 'abductive inference', 'abductive logic' and 'abductive theory of method'. A commonly used synonym with a different root word was 'retroduct 'or 'retroduction'. Moreover, other uncommon surrogate terms were also used which aimed to convey the same meaning as abductive reasoning. These included: 'synthetic reasoning', 'fuzzy logic', 'if/then/therefore reasoning'.

As a synonym of abductive reasoning, synthetic reasoning allows data to be synthesized to figure out what one ought to do and, as a result, leads to the formation of a theory or hypothesis (Raholm 2010a, Raholm 2010b). The term fuzzy logic is a form of vague, imprecise and probabilistic reasoning (Rolfe 1997), which is similar to the if/then/therefore way of thinking concerned with plausible hypothesis generation. Retroduction, which has been used interchangeably with abductive reasoning, is an inference through which events are explained by suggesting the existence of mechanisms which could have produced them (Sayer 1992). This is similar to the definition of abductive reasoning which has been declared earlier as synthetic reasoning which focuses on underlying patterns of phenomena to generate plausible hypotheses.

Related Terms or Concepts

Terms such as ‘data-driven reasoning’, ‘forward reasoning’ and ‘bottom-up reasoning’ are similar to abductive reasoning. These forms of reasoning aim to generate a hypothesis or explanation from observations or data (i.e., data-to-hypothesis or data-to-explanation). This is similar to inductive reasoning where the explanation or hypothesis is a confirmation of the presented data and not a plausible inference from the underlying patterns of certain phenomena to which the data are hinting. However, terms such as deduction, induction, clinical reasoning and their synonyms were not included in the list of related terms because they have been discussed earlier as background information and straying into a lengthy discussion about them goes beyond the scope of this concept analysis.

Other terms related to abductive reasoning included: ‘exploratory factor analysis’, ‘analogical modeling/reasoning/transfer/inference’, ‘similarity recognition’ and ‘intuition’. ‘Exploratory factor analysis’ and ‘analogical modeling/reasoning/transfer/inference’ were specific to the psychology and medical literature. These two terms were neither explained nor elaborated on in the sources, which could be due their common use in the field of statistics. While the former is involved in the identification of underlying relationships between variables (Norris & Lecavalier 2010), the latter is a way of using analogy to represent a certain phenomenon in the world so the problem can be examined in new ways and with newer ideas (Truitt & Rogers 1960).

Similarity recognition refers to the unconscious matching of a presenting clinical situation to a similar, previously encountered one (Brooks *et al.* 1991, Norman & Brooks 1997). Rolfe (1997) critiques Benner (1984)'s work and states that she was unable to explain expertise and labelled it as intuition. Intuition, as it is understood, fails to allow a nurse to verbalize and justify professional judgments and clinical decisions as it is based on some form of instinctive cognitive process that lacks conscious reasoning and is often based on similarity or pattern recognition. In contrast, abductive reasoning allows a nurse to systematically explain and justify how he came to his clinical decisions (Raholm 2010a, Rolfe 1997).

Antecedents and Consequences

Antecedents of abductive reasoning were divided into two domains. The first domain was related to the clinical situation presented before a clinician, which warranted explanation. This domain included terms such as: 'complex patterns of reality' and several combinations of 'surprising or puzzling facts', 'data', 'experience', 'phenomena' and/or 'symptoms'. The second domain was related to the characteristics a clinician must possess, successfully, to detect, extract and draw meaning from the surprising or puzzling clinical situation. Terms included in this domain were: 'creativity'; 'boldness'; 'similar previous experience'; and 'strong knowledge' that were 'tacit', 'pre-existing', 'theoretical', 'established', 'declarative', 'conceptual', 'experiential' and/or 'empirical'.

Consequences of abductive reasoning were also divided into two categories. The first category suggested endpoints that would result from engaging in the process of abductive reasoning. This category contained specific items such as: 'deeper knowledge and/or understanding'; 'narratives'; 'breadth'; 'cases'; and 'hypotheses' or 'explanations' which were 'explanatory', 'diagnostic' and/or 'plausible'. These narratives and cases could be described as hypothetical explanations which are formulated to describe a breadth of nursing issues in a care situation. For example, after thorough assessment of a newly admitted client, a nurse may write a hypothetical case describing his interaction with the patient and nursing-related issues which will guide further care planning. The second category emphasized cognitive processes such as: 'deduction', 'induction' and 'scientific progress'. This shows that once hypotheses are formulated through abduction, they are explicated logically through deduction and then empirically through induction (Eriksson & Lindström 1997).

Defining Attributes

Several defining attributes of abductive reasoning surfaced in the concept analysis. First, abductive reasoning was defined by characteristics which included: 'uncertainty'; 'educated guess'; and 'fuzzy reasoning'. These terms emphasize that abductive reasoning generates plausible explanations which are vague and imprecise (i.e., fuzzy), could be considered educated guesses and have an element of uncertainty. Secondly, abductive reasoning was described as: 'visual abduction', 'creative abduction and selective abduction'; while other defining attributes included 'inference to the best explanation' and 'explanatory coherence'.

According to Magnani (1997), visual abduction is based on similarity recognition; creative abduction is an overarching term that deals with the field concerned with scientific growth; while selective abduction allows for a diagnostic hypothesis to be selected from a pre-specified set of hypotheses that have been established. Since abductive reasoning involves two types of reasoning approaches: (1) generation of hypothesis; and (2) evaluation of hypothesis, Haig (2008) distinguishes between the two by explaining hypothesis generation as such (p. 1014):

1. The surprising empirical fact F is detected.
2. But if hypothesis H were approximately true, then F would follow as a matter of course.
3. Hence, there is reason to believe that H is plausible.

The second form of abductive reasoning is referred to as inference to the best explanation, which allows for the evaluation of competing explanatory hypotheses to select the best among them. This form of inference is centrally concerned with the establishment of explanatory coherence which argues that a theory's propositions remain unified mainly because of their explanatory relations (Thagard 1978). Explanatory coherence is used to evaluate explanatory theories or hypotheses and consists of three criteria: (a) 'explanatory breadth' determines whether the hypothesis explains the greatest range of facts which cannot be explained by rival explanations; (b) 'simplicity' ensures that the hypothesis is a simple explanation with the fewest assumptions and (c) 'analogy' is used as a credible explanation to support the hypothesis (Thagard 1978). Haig (2008) describes inference to the best explanation in the following general schema (p. 1015):

1. F1, F2, … are surprising empirical facts.
2. Hypothesis H explains F1, F2, …
3. No other hypothesis can explain F1, F2, … as well as H does.
4. Therefore, H is accepted as the best explanation.

Abductive reasoning has also been described as the process of 'hypothesis or theory generation/creation'. The process of generating a hypothesis in the context of clinical reasoning has been explained by Vertue and Haig (2008)'s a five-step process (Table A1.3) which is based on Haig (2005)'s Abductive Theory of Method. The first phase, 'phenomena detection', involves both data collection and analysis. Several data collection strategies are utilized to collect high quality data associated with the phenomena being investigated. Data are then analyzed according to pre-existing knowledge, experience, etc., to recognize emerging patterns that evince the presence of several phenomena. In the second phase, 'inferring causal mechanisms', the clinician uses a framework (e.g., biopsychosocial model) to identify and group relevant plausible causal factors and suggest their potential relationship to the detected phenomena (Vertue & Haig 2008).

The third phase, 'developing a causal model', allows a clinician to develop a model where the relationships and interactions of the various causal mechanisms are considered. Core mechanisms that are more centrally involved in causal relationships are often the ones that require immediate intervention. Once all causal relationships are described, the fourth phase, 'evaluating the causal model', allows the clinician to evaluate the causal model by determining its explanatory coherence. Following this, the clinician conceptualizes the clinical situation in the most explanatorily coherent way. The final phase, 'formulating the case', marks the conclusion of the abductive clinical reasoning process. The formulated case is a comprehensive and integrated conceptualization of descriptive and explanatory hypotheses. It emphasizes the possible links between the various aspects of a care situation (Vertue & Haig 2008).

Table A1.3: Abductive Reasoning Process*

Phenomena Detection	Inferring Causal Mechanism	Developing a Causal Model	Evaluating the Causal Model	Formulating the Case
• Clinician collects and analyzes cues from the presenting situation.	• Clinician identifies and groups relevant plausible causal factors and suggests their possible relationship to the detected phenomena. • Clinician uses biopsychosocial framework to structure his or her thinking process.	• Clinician develops an illustration where various causal mechanisms are considered.	• Clinician ensures all relationships are coherent and supported by data. • Clinician conceptualizes the care situation in the most explanatorily coherent way.	• Clinician emphasizes the possible links between the various aspects of a care situation through comprehensive and integrated conceptualization of descriptive and explanatory hypotheses.

*Based on Vertue & Haig (2008)

Examples and Model Case

While an article related abductive reasoning to jury decision-making and crime detection (Haig 1999), another related it to detectives who search for anomalous phenomena (Raholm 2010a). Both examples, however, did not elaborate nor clearly explained how abductive reasoning was used. Another paper explained how computers flew helicopters because abductive reasoning allowed them to learn fuzzy rules which could be justified and verbalized to show patterns of thinking that resembled those of experts (Rolfe 1997). An example based on the use of the hypothetico-deductive method of reasoning, showed how a clinician recommended surgery after erroneously deducing from a client's blood results, ultrasound and abdominal inspection that gallbladder stones were the cause of her pain symptoms without entertaining other tentative hypotheses (Lawson & Daniel 2011).

Table A1.4: Model Case of Abductive Reasoning

	Deduction	Induction	Abduction
Niiniluoto (1999)	• All the beans from this bag are white. • These beans are from this bag. • These beans are white.	• These beans are from this bag. • These beans are white. • All the beans from this bag are white.	• All the beans from this bag are white. • These beans are white. • *These beans are from this bag.
Magnani (1997)	• If a patient is affected by a beta-thalassemia, his level of hemoglobin A2 is increased. • John is affected by a beta-thalassemia. • John's level of hemoglobin A2 is increased.	• John is affected by a beta-thalassemia. • John's level of hemoglobin A2 is increased. • If a patient is affected by a beta-thalassemia, his level of hemoglobin A2 is increased.	• If a patient is affected by a beta-thalassemia, his level of hemoglobin A2 is increased. • John's level of hemoglobin A2 is increased. • *John is affected by a beta-thalassemia.
Proposed nursing-focused model case	• If a patient takes calcium, she could have constipation. • Mrs. Smith takes calcium. • Mrs. Smith could be constipated.	• Mrs. Smith takes calcium. • Mrs. Smith is constipated. • If a patient takes calcium, she could have constipation.	• If a patient takes calcium, she could have constipation. • Mrs. Smith is constipated. • *Mrs. Smith takes calcium.

*Plausible abductive inferences

While other articles did not present daily-life examples of abductive reasoning, they did however present general schema of the process of abductive reasoning compared with the process of inductive and deductive reasoning. Since Rodgers (1989) recommends that a model case to validate the concept and its characteristics be identified rather than constructed, a general schema (Niiniluoto 1999) and a similar medical exemplar schema (Magnani 1997) were identified and presented in Table A1.4. Building on these examples, a nursing-focused schema was also proposed (Table A1.4) as an additional model case to show a practice example of how abductive reasoning could be used in everyday clinical nursing practice.

Discussion

Analysis Method and Data Sources

The concept analysis of abductive reasoning followed Rodgers (1989)'s evolutionary approach to concept analysis. This method was easy to use and apply. It was also suitable for abductive reasoning because abductive reasoning is an evolving concept. Although only twelve articles were used for the concept analysis, saturation was achieved within and between disciplines regarding what abductive reasoning is and is not as a concept. Of the three articles found in medicine, Lawson and Daniel (2011)'s and Magnani (1997)'s articles were specific to medical practice and provided concrete examples; while Upshur (1997)'s paper was only partially relevant and consisted of vague conclusions, of which one was an incorrect abductive inference that was presented as an example along with the general schema.

All articles from the discipline of psychology were very useful in highlighting the importance of abductive reasoning in clinical reasoning and practice. However, all of these articles were either authored or co-authored by Brian Haig, whose five-step process of abductive reasoning (Vertue & Haig 2008) has been described in the analysis and is a useful method which can also be applied to nursing practice. Nursing articles on abductive reasoning were primarily philosophy-based and lacked concrete nursing practice examples to highlight how abductive reasoning could be implemented into nursing practice or nursing education. This was mainly due to the context of four of the five nursing articles which focused on knowledge or theory development and not on clinical practice. The remaining paper (Rolfe 1997) examined clinical expertise but compared nurses to computers.

Nursing Implications

Novice nurses readily retrieve clinical data, but often overlook important cues when confronted with increasing complexity of situation and heightened degree of uncertainty (Andersson *et al.* 2006; O'Neill *et al.* 2005). This is mainly because nursing curricula immerse nursing students in either the hypothetico-deductive method of reasoning (Wong & Chung 2002), or both inductive and deductive approaches to reasoning (Chinn & Kramer 2011). It has been discussed that while the hypothetico-deductive method develops and tests specific testable hypotheses based on limited data, induction focuses on forming generalizations. As an alternative approach, the holistic approach of abductive reasoning can allow nursing students (and students of other health disciplines) to build hypotheses through maximum data retrieval and to develop causal models which illustrate and explain the underling structures of the situation.

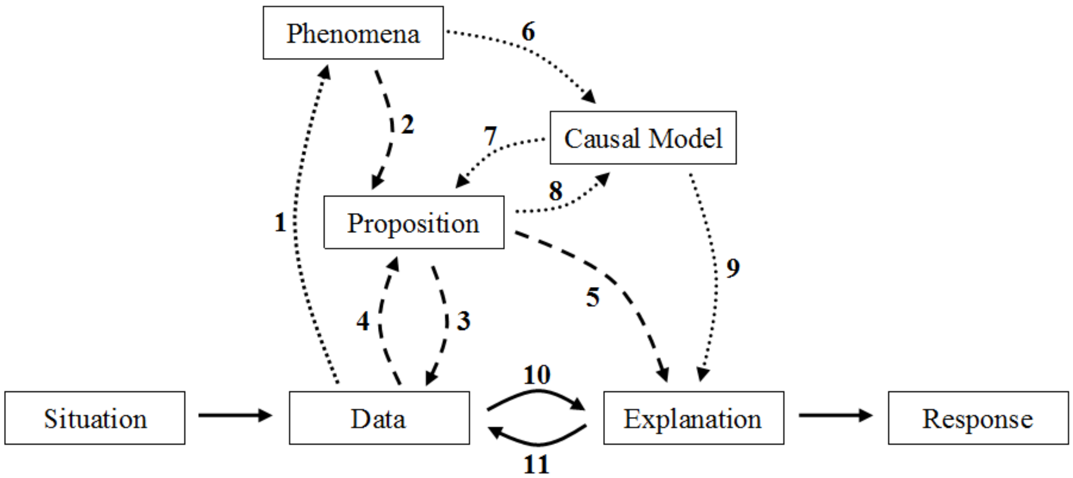
Many prominent experts in the area of knowledge development in nursing recognize only inductive and deductive modes of reasoning (Chinn & Kramer 2011). These experts further discuss various patterns or ways of knowing which include: empirical, ethical, aesthetic, personal and emancipatory knowing. It can be argued that inductive and deductive modes of reasoning are limited in their scope and cannot cater to the various ways of knowing altogether. While the hypothetico-deductive (which includes deductive) approach comes from an empirical paradigm, it mainly favours empirical knowing while the inductive approach which comes from an interpretive paradigm favours non-empirical forms of knowing (Monti & Tingen 1999).

Arising from pragmatism, abductive reasoning appears to favour both empirical and non-empirical forms of knowing. This is illustrated in this concept analysis where both subjective and objective data must be entertained and analyzed to recognize phenomena. This describes ‘personal’, ‘empirical’ and ‘aesthetic’ ways of knowing where empirical is the objective data, aesthetic is the subjective data and personal is the interpreter’s analysis of these data. Furthermore, Raholm (2010a) points out that abductive reasoning is related to what nurses ought to do which relates to nurses’ ethical obligation and points towards ‘ethical’ knowing. Although ‘emancipatory’ knowing is not discussed in the limited literature explored for this concept analysis, it can be proposed that this form of knowing is recognized during the causal model phase of abductive reasoning where the underlying patterns of certain phenomena are explored and understood in light of the presenting situation.

In problem-based learning (PBL) which is founded on the hypothetico-deductive approach (Barrows & Tamblyn 1980) and is commonly used in several nursing programs (Rideout & Carpio 2001), abductive reasoning can offer an alternative approach to learning which not only focuses on hypothesis-building, but also incorporates different ways of knowing to encourage students to examine situations from various perspectives. While the hypothetico-deductive approach uses limited data to generate early hypotheses which are tested through further data collection, the abductive reasoning approach examines all sorts of data through the use of different ways of knowing and then discovers phenomena and their possible causes which are interlinked with one another to generate a plausible explanation that explains all or most of the issues in a presenting situation. Hence, in PBL curricula where the emphasis is to explore situations and explain them through learning and knowledge-building, abductive reasoning may be a more appropriate choice than the hypothetico-deductive method which aims to test hypotheses and resolve situations based on limited and goal-oriented data collection.

Reasoning Model

Proposed in Figure A1.2 is a visual model which attempts to show different types of reasoning approaches discussed in the analysis. After confrontation with a ‘situation’, initial ‘data’ from the situation are obtained. According to the hypothetico-deductive method, a clinician recognizes cues, articulates early hypotheses or propositions, collects more data either to confirm or refute his initial propositions and formulates an explanation based on the accepted propositions to ensure that their consequences are 'explained' logically. This hypothetico-deductive process is shown by arrows 1 to 5.



1. Phenomena detection, cue recognition
2. Preliminary hypothesis retrieval
3. Data collection
4. Proposition confirmation, cue interpretation
5. Hypothesis evaluation, diagnosis
6. Inference of causal mechanisms
7. Proposition creation/generation
8. Coherent conceptualization of care situation
9. Case formulation, inference to the best explanation
10. Induction, similarity/pattern recognition, intuition, data-driven reasoning
11. Deduction, hypothesis-driven reasoning

Figure A1.2. Proposed model of reasoning

In abductive reasoning, a clinician detects cues or 'phenomena' and infers their causal mechanisms to develop a visual 'causal model'. In this illustration, various causal mechanisms and 'propositions' are considered and evaluated to generate a coherent 'explanation'. This abductive reasoning process is represented by arrows 1, 6, 7, 8 and 9. Inductive reasoning, demonstrated by arrow 10, allows a clinician to formulate an *explanation* through intuition or pattern recognition where the clinician automatically links the present situation to a similar situation which he has learned about or witnessed in the past. Arrow 11 displays deductive reasoning which confirms specific 'data' from general 'explanations' which are known.

Limitations

There were several limitations in this concept analysis. Data sources that were used for the analysis were limited to electronic databases only. The focus of the data sources was also only on three disciplines while literature from other disciplines that also engage in clinical reasoning were not included such as: physiotherapy, occupational therapy, speech language pathology, dietetics, etc. Although these disciplines were not included, there was no intention to exclude them either. Should publications focused on abductive reasoning and clinical practice had emerged from their disciplinary bodies of literature, they would have been included. However, in the databases used, key literature from only three disciplines surfaced.

A major limitation in this concept analysis was that clinical reasoning as a key word was used, while its surrogate terms were left out such as decision making, problem solving, diagnostic reasoning and clinical judgment. This could have opened up other branches of literature, perhaps even literature from other health disciplines. What is important is that saturation was achieved. However, since concepts are always evolving, the concept of abductive reasoning may continue to change in the future as newer literature emerges that examines it in the context of clinical reasoning and clinical practice, or another context unknown or unfamiliar to us now.

Conclusion

It is important for nurses to be aware of the reasoning strategies they utilize when engaged in clinical practice. Abductive reasoning can allow nurses to recognize the deeper underlying patterns or phenomena in complex clinical situations (Eriksson & Lindström 1997). It is predicted that novice nurses and other clinicians can begin to use expert approaches to reasoning if they are taught how to think like experts (Rolfe 1997). Vertue and Haig (2008) offer a model that highlights the various phases of abductive reasoning which nurses and other clinicians can follow to develop abductive reasoning skills and consequently improve their clinical reasoning abilities. These steps, along with steps of inductive and hypothetico-deductive reasoning, are also outlined in the proposed reasoning model in Figure A1.2.

Training nursing students or practicing nurses in abductive reasoning could help them develop hypothesis-building skills which can enhance their reasoning abilities both in educational and clinical contexts. However, research is needed to determine how attributes of abductive reasoning can be strengthened. This can include both quantitative and qualitative research approaches that aim to develop and test specific educational interventions that target certain attributes of abductive reasoning such as the breadth and depth of phenomena detection; accuracy, relevancy and coherence of generated causal mechanisms and their relationships; the breadth and depth of causal models; and the accuracy of case conceptualizations in explaining a particular situation.

References

Andersson A, Omber M & Svedlund M (2006) Triage in the emergency department – a qualitative study of the factors which nurses consider when making decisions. *Nursing in Critical Care* 11, 136-145.

Barrows HS & Tamblyn RM (1980) *Problem-based learning: an approach to medical education.* Springer Publishing Company, New York.

Benner P. (1984) *From novice to expert: excellence and power in clinical nursing practice.* Addison-Wesley, Menlo Park CA.

Brooks LR, Norman GR & Allen SW (1991) Role of specific similarity in a medical diagnostic task. *Journal of Experimental psychology: General* 120, 278-287.

Chinn PL & Jacobs JK (1983) *Theory and nursing: a systematic approach*. C.V. Mosby, St. Louis MO.

Chinn PL & Kramer MK (2011) *Integrated theory and knowledge development in nursing* (8th ed.). Mosby, St. Louis MO.

Coderre S, Mandin H, Harasym PH & Fick GH (2003) Diagnostic reasoning strategies and diagnostic success. *Medical Education* 37, 695-703.

Elstein AS, Shulman LS & Sprafka SA (1978) *Medical problem solving: an analysis of clinical reasoning.* Harvard University Press, Cambridge MA.

Eriksson K & Lindström UA (1997) Abduction – a way to deeper understanding of the world of caring. *Scandinavian Journal of Caring Sciences* 11, 195-198.

Fann, K.T. (1970). *Peirce's theory of abduction.* The Hague, Netherlands: Martinus Nijhoff.

Gordon M (1987) *Nursing diagnosis: process and application*. McGraw-Hill, New York.

Haig BD (1999) Construct validation and clinical assessment. *Behaviour Change* 16, 64-73.

Haig BD (2005) An abductive theory of scientific method. *Psychological Methods* 10, 371-388.

Haig BD (2008) Scientific method, abduction and clinical reasoning. *Journal of Clinical Psychology* 64, 1013-1018.

Lawson AE & Daniel ES (2011) Inferences of clinical diagnostic reasoning and diagnostic error. *Journal of Biomedical Informatics* 44, 402-412.

Magnani L (1997) Basic science reasoning and clinical reasoning intertwined: epistemological analysis and consequences for medical education. *Advances in Health Sciences Education* 2, 115-130.

McFadden EA & Gunnett AE (1992) A study of diagnostic reasoning in pediatric nurses *Pediatric Nursing* 18, 517-538.

Monti EJ & Tingen MS (1999) Multiple paradigms of nursing science. *Advances in Nursing Science* 21, 64-80.

Newell A & Simon HA (1972) *Human problem solving*. Prentice-Hall, Englewood Cliffs NJ.

Niiniluoto I (1999) Defending abduction. *Philosophy of Science* 66, S436-S451.

Norman GR & Brooks LR (1997) The non-analytical basis of clinical reasoning. *Advances in Health Sciences Education* 2, 173-184.

Norris M & Lecavalier L (2010) Evaluating the use of exploratory factor analysis in developmental disability psychological research. *Journal of Autism and Developmental Disorders* 40, 8-20.

O'Neill E, Dluhy NM & Chin E (2005) Modelling novice clinical reasoning for a computerized decision support system. *Journal of Advanced Nursing* 49, 68-77.

Patel VL, Arocha JF & Zhang J (2005) Thinking and reasoning in medicine. In *The Cambridge handbook of thinking and reasoning* (Holyoak KJ & Morrison RG eds.), Cambridge University Press, New York, pp.727-750.

Peirce, C.S. (1903 reprinted 1998). Pragmatism as the logic of abduction. In the Peirce Edition Project (Ed.), *The essential Peirce: Selected philosophical writings: Volume 2 (1893-1913)* (pp. 226-241). Bloomington, IN: Indiana University Press.

Peirce, C.S. (1931-1958). Collected papers of Charles Sanders Peirce. Vol. 1-6, C. Hartshorne, & P. Weiss (Eds.). Vol. 7-8, A. W. Burks (Ed.). Cambridge, MA: Harward University.

Radwin LE (1995) Conceptualizations of decision making in nursing: analytic models of 'knowing the patient.' *Nursing Diagnosis, 6*(1), 16-22.

Raholm MB (2010a) Abductive reasoning and the formation of scientific knowledge within nursing research. *Nursing Philosophy* 11, 260-270.

Raholm MB (2010b) Theory development and the logic of discovery. *International Journal of Human Caring* 14, 36-42.

Reed GP (1995) A treatise on nursing knowledge development for the 21st century: beyond postmodernism. *Advances in Nursing Science* 17, 70-84.

Rideout E & Carpio B (2001) The problem-based learning model of nursing education. In *Transforming nursing education through problem-based learning* (Rideout E ed.), Jones and Bartlett Publishers, Sudbury MA, pp. 21-49.

Rimoldi HJA (1961) The test of diagnostic skills. *Journal of Medical Education* 36,73-79.

Rodgers B (1989) Concepts, analysis and the development of nursing knowledge: the evolutionary cycle. *Journal of Advanced Nursing* 14, 330-335.

Rodgers B (2000) Philosophical foundations of concept development. In *Concept development in nursing: foundations, techniques and applications* (Rodgers B & Knafl K eds.), W.B. Saunders, Philadelphia PA, pp. 7-38.

Rolfe G (1997) Science, abduction and the fuzzy nurse: an exploration of expertise. *Journal of Advanced Nursing* 25, 1070-1075.

Sayer A (1992) *Method in social science: a realist approach* (2nd edn.). Routledge, London.

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

Smith EE & Medin DL (1981) *Categories and concepts*. Harvard University Press, Cambridge MA.

Tanner CA (1982) Instruction on the diagnostic process: an experimental study. In M.J. Kim & D.A. Mortiz eds.), *Classification of nursing diagnosis* (Kim MJ & Mortiz DA eds.), McGraw-Hill, New York, pp. 145-152).

Tanner CA, Padrick KP, Wetfall UA & Putzier DJ (1987) Diagnostic reasoning strategies of nurses and nursing students. *Nursing Research* 36, 358-363.

Thagard P (1978) The best explanation: Criteria for theory choice. *Journal of Philosophy* 75, 76-92.

Tofthagen R & Fagerstrøm LM (2010) Rodgers' evolutionary concept analysis – a valid method for developing knowledge in nursing science. *Scandinavian Journal of Caring* 24, 21-31.

Truitt TD & Rogers AE (1960) *Basics of analog computers*. John F. Rider Publishing, New York.

Upshur R (1997) Certainty, probability and abduction: Why we should look to C.S. Peirce rather than Godel for a theory of clinical reasoning. *Journal of Evaluation in Clinical Practice* 3, 201-206.

Vertue FM & Haig BD (2008) An abductive perspective on clinical reasoning and case formulation. *Journal of Clinical Psychology* 49, 1046-1068.

Walker LO & Avant KC 1983. *Strategies for theory construction in nursing.* Appleton-Century-Crofts, Norwalk CT.

Ward T & Haig B (1997) Abductive reasoning and clinical assessment. *Australian Psychologist* 32, 93-100.

White JE, Nativio DG, Kobert SN & Engberg SJ (1992) Content and process in clinical decision-making by nurse practitioners. *Journal of Nursing Scholarship* 24, 153-158.

Wong TKS & Chung JWY (2002) Diagnostic reasoning processes using patient simulation in different learning environments. *Journal of Clinical Nursing* 11, 65-72.

Appendix 2: Comparative Analysis of External Validity Reporting in Non-randomized Intervention Studies

CJNR
School of Nursing
McGill University

3506 University Street
Montreal, Quebec, Canada H3A 2A7
(514) 398-4160 Fax (514) 398-8455
cjnr.nursing@mcgill.ca
www.cjnr.nursing.mcgill.ca

COPYRIGHT PERMISSION

DATE:	March 18, 2015
TO:	Noeman Mirza McMaster University School of Nursing - Graduate Studies 1280 Main St. West HSC 3H48 Hamilton, ON L8S4K1
ARTICLE:	Mirza, N.A., Akhtar-Danesh, N., Staples, E., Martin, L., Noesgaard, C. (2014). Comparative analysis of external validity reporting in non-randomized intervention studies.*CJNR (Canadian Journal of Nursing Research) 46*(4), 47-64.
PERMISSION FOR:	The sole usage by author Mirza to include the above mentioned article as an appendix to his PhD thesis. Deadline date December 31, 2015.

Please return a copy of this invoice with your payment. MC and Visa credit cards are accepted. Check or money order is payable to the *Canadian Journal of Nursing Research.*

Thank You.

Joanna Toti
Managing Editor, *CJNR*

Abstract

PURPOSE: To describe a comparative analysis of external validity reporting in non-randomized behavioural and public health intervention studies that used and did not use the TREND (Transparent Reporting of Evaluations with Non-randomized Designs) statement.

DESIGN: Comparative analysis.

FINDINGS: The search resulted in fourteen non-randomized intervention studies which were rated based on Green and Glasgow (2006)'s criteria for external validity reporting. Studies that used the TREND statement demonstrated improved external validity reporting when compared to studies that did not use the TREND statement.

IMPLICATIONS: The TREND statement and Green and Glasgow's criteria can help improve external validity reporting of non-randomized behavioural and public health interventions.

KEYWORDS: TREND statement, external validity, comparative analysis, public health, non-randomized interventions

Introduction

Over the past two decades, several guidelines have emerged which aim to enhance the quality of reports of randomized controlled trials (RCTs), non-randomized experiments, systematic reviews, and meta-analyses. The TREND (Transparent Reporting of Evaluations with Non-randomized Designs) statement is used to enhance the quality of reporting in non-randomized intervention studies. However, its impact on external validity reporting is unclear. Therefore, this comparative analysis aims to determine whether the use of the TREND statement enhances external validity reporting in non-randomized intervention studies. To do this, both TREND and non-TREND studies are evaluated by external validity criteria recommended by Green and Glasgow (2006). Findings and implications for nurse researchers who are engaged in conducting, reporting, and evaluating studies involving non-randomized interventions are discussed.

Background

Investigators concerned with health promotion engage in clinical research in order to draw inferences from study findings about the nature of their surroundings. To interpret study findings, two sets of inferences are commonly used. The first inference, known as *internal validity*, is the extent to which correct conclusions are drawn about what actually happened in an experiment; while the second inference, *external validity* (i.e., *generalizability*), is the extent to which study findings can be applied to situations outside the experiment (Hulley, Cummings, Browner, Grady, & Newman, 2006). For an accurate interpretation of study findings to occur, a study must first have strong internal validity which is achieved through a strong relationship between its research operations built upon good choice of study design, outcome measurement, and representative sampling. It is for this reason that researchers and journals give precedence to internal validity and scientific rigor instead of generalizability of study findings (Ferguson, 2004). This jeopardizes translation of research into practice of applied disciplines such as medicine, public health and nursing, which are concerned about health promotion and improving the health of the public (Steckler & McLeroy, 2008).

Balas and Boren (2000) claimed that it takes several years to translate even small amounts of original research into interventions that enhance patient care. They attributed this delay partly to the inadequacy of how healthcare providers are assisted in assessing the strengths of study results and applying them to practice. Over the last decade, after the introduction of the CONSORT (Consolidated Standards of Reporting Trials) statement which aims to improve quality of reporting of RCTs (Begg et al., 1996), focus on the methodological quality of research reports has increased (Moher et al., 2010). However, reporting criteria of the CONSORT statement emphasize internal validity while they do not address external validity in its entirety (Glasgow et al., 2006). Reviews show that lack of discussion on external validity disadvantages judgment around the potential effectiveness of interventions and their applicability to practice (Glasgow, Klesges, Dzewaltowski, Bull, & Estabrooks, 2004). Therefore, there is a need to strengthen the reporting of generalizability of research findings (Ferguson, 2004).

Given that RCTs are not always feasible and may not be ethical within public health (Victora, Habicht, & Bryce, 2004), the TREND statement was developed to improve the quality of reporting of non-randomized evaluations of behavioral and public health interventions (Des Jarlais, Lyles, Crepaz, & the TREND Group, 2004). After its publication, the statement drew immediate praise from editors of several journals (Caetano, 2004; Kirkwood, 2004; Ross, Elford, Sherr, & Hart, 2004; Treasure, 2004). However, it was also criticized for its limited external validity criteria which were viewed as insufficient for reporting and evaluating the generalizability of study results (Dzewaltowski, Estabrooks, Klesges, & Glasgow, 2004). These critics insisted on additional criteria related to external validity. Green and Glasgow (2006) later elaborated on this concern by proposing criteria for external validity reporting (Table A2.1).

Purpose

Since its introduction in 2004, several researchers have used the TREND statement as a guideline for reporting studies involving non-randomized designs. To the best of our knowledge, the impact of the use of TREND statement guidelines on external validity reporting of non-randomized intervention studies has not previously been reported in the literature. Therefore, the purpose of this comparative analysis was to fill this gap with objectives which were threefold:

1. Review selected reports claiming to have used the TREND statement as a guideline (i.e., TREND studies) and evaluate the extent to which these studies report external validity;
2. Review selected recent reports that did not use the TREND statement as a guideline (i.e., non-TREND studies) and evaluate the extent to which these studies report external validity; and

3. Offer a comparative overview of external validity reporting of both TREND and non-TREND studies.

Literature Search

TREND Studies

Before the literature was searched, it was decided that the analysis would focus on prospective non-randomized intervention studies with a comparison group and a follow-up. No date limitations were set when searching for TREND studies because the TREND guideline was published in 2004. To do this, the original TREND article by Des Jarlais et al. (2004) was sought in several databases (Figure A2.1), after which its citations (i.e., articles citing the original TREND article) in each database were examined. The combined search resulted in 558 records, from which 515 records were excluded because these were not intervention studies. Of the remaining 43 records, 28 were excluded because they were either duplicates or cited the TREND guidelines without mentioning, discussing or declaring whether or not the TREND guidelines were used for reporting. This resulted in 15 records, of which 8 were excluded because they were either study protocols, used retrospective study designs, had no comparison group, lacked follow-up, or referred to their post-test as follow-up. This generated a total of 7 TREND studies which were included in the analysis.

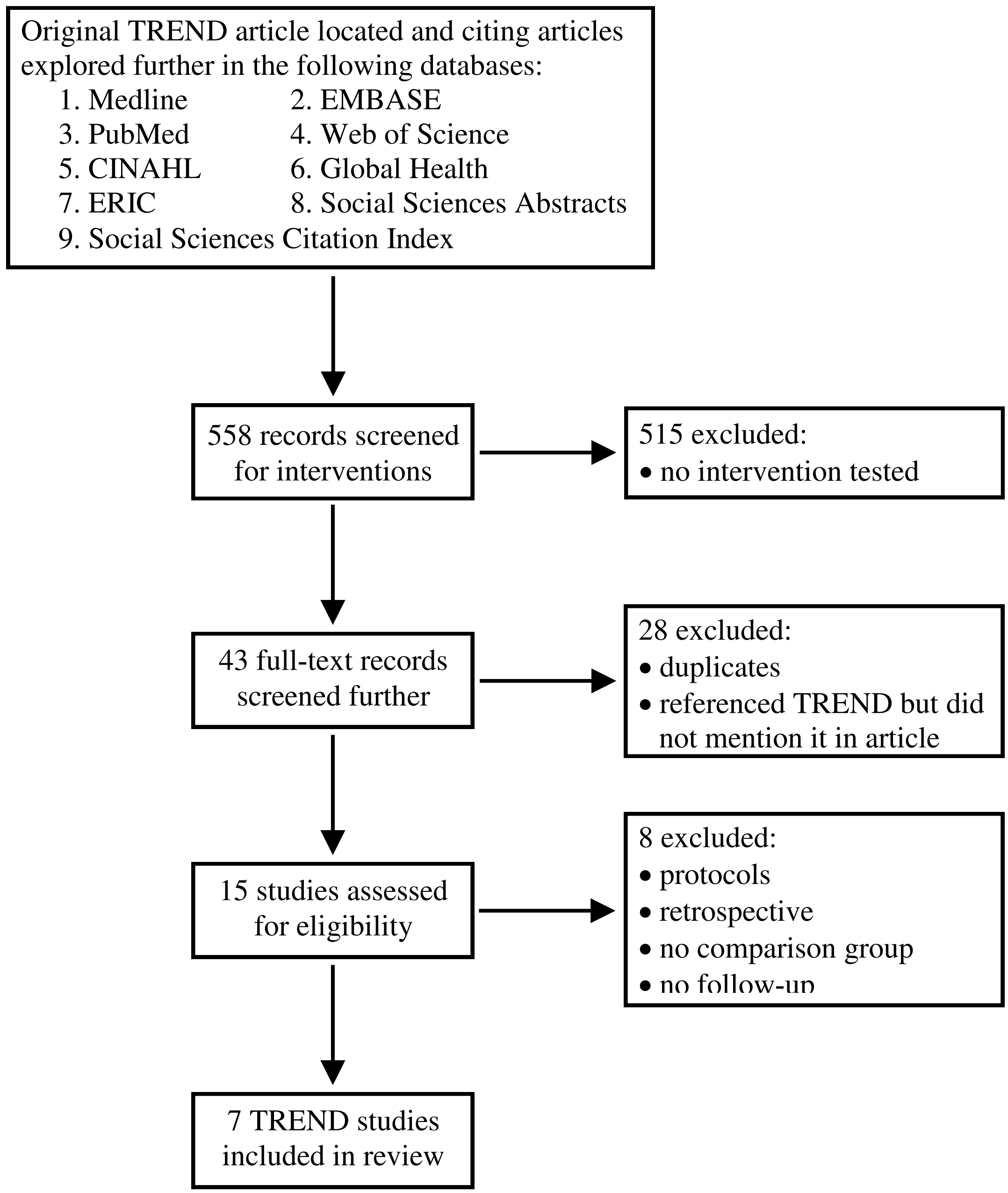


Figure A2.1. Search strategy for TREND studies

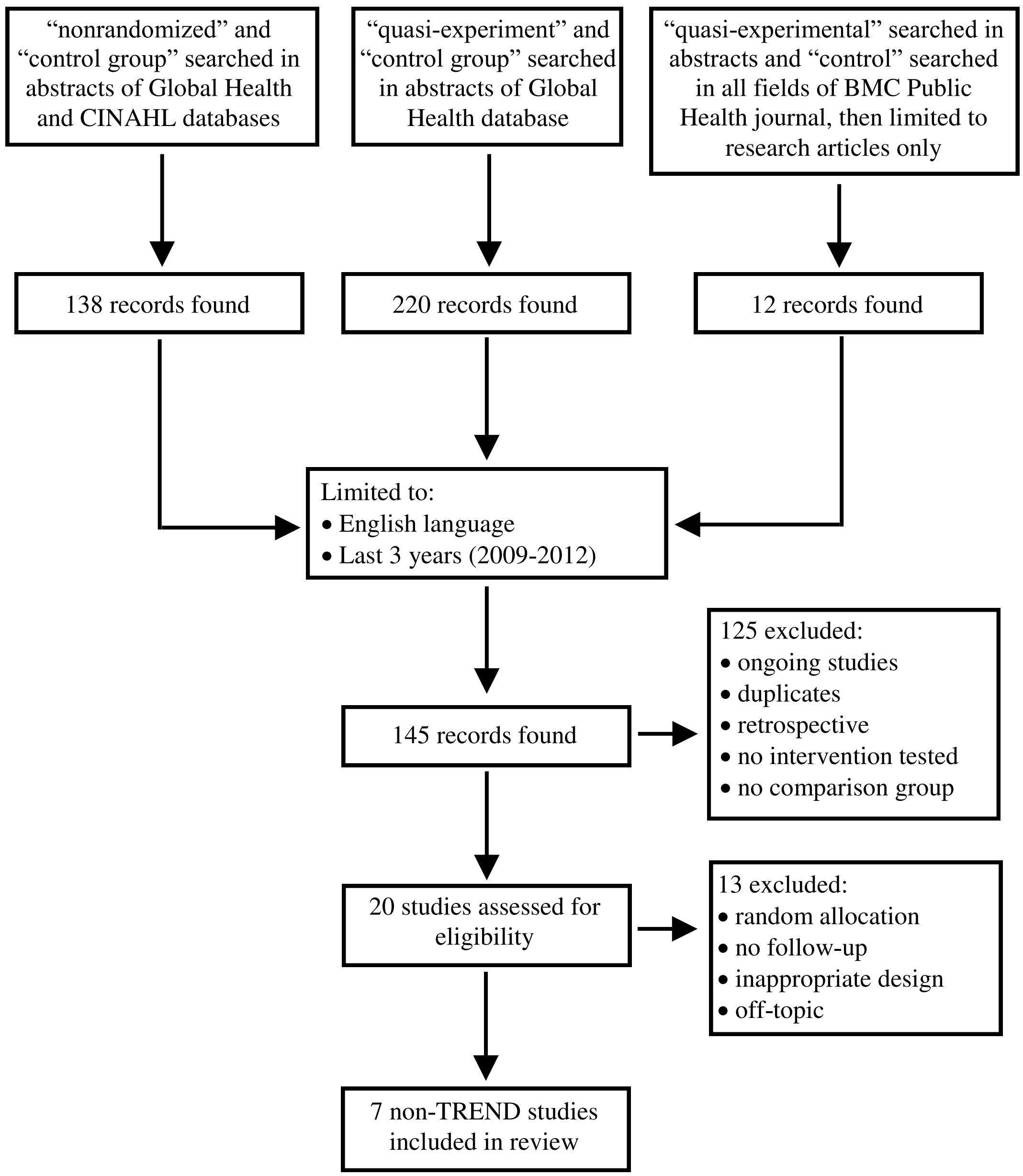


Figure A2.2. Search strategy for non-TREND studies

Non-TREND Studies

To compare the 7 TREND studies with non-TREND studies, 7 non-TREND studies were sought. Selection criteria for non-TREND studies were similar to those of TREND studies (i.e., non-randomized intervention studies with control group and follow-up). While the search for TREND studies permitted studies to date back to 2004 (when TREND guidelines were published), only recent non-TREND studies were obtained. This was done by examining the most recent studies first and then moving back in time until 7 non-TREND studies were obtained. Reasons to use recent studies were based on the assumption that reporting of research studies should improve over time and recent study reports would represent improved trends in reporting. Secondly, it was assumed that the five-year gap from 2004 (when TREND guidelines were published) to 2009 was sufficient to allow the uptake of such guidelines by the research community. Therefore, recent studies were limited to those published in 2009 or after.

Since the TREND statement was developed initially for behavioural and public health interventions, popular nursing and public health databases were selected (i.e., Global Health, CINAHL). Key terms such as *nonrandomized* and *control group* were used in study abstracts of CINAHL and Global Health databases (Figure A2.2). This resulted in 138 records. The term *quasi-experiment* was also used in the Global Health database, which generated 220 records. Furthermore, the term *quasi-experiment* was also used in the BMC Public Health journal since this was a common journal among the selected TREND studies. This search generated 12 records. All 370 records were limited to the English language and to as early as 2009. This resulted in 145 records, from which

125 were excluded because they were ongoing studies (incomplete), used retrospective design, did not test an intervention (e.g., survey), had no comparison group, or were duplicate records. From the 20 remaining records, 13 were further excluded because they had randomized allocation, lacked follow-up, had an inappropriate design (e.g., mentioned quasi-experiment but were cross-sectional studies), or strayed from the theme of behavioural and public health interventions. Coincidentally, this search also resulted in 7 non-TREND studies which ranged from 2009 to 2011. If there had been more or less 7 non-TREND studies, the year of publication would have been adjusted to either 2010 or 2008 respectively in order to have a comparable number of TREND and non-TREND studies for the analysis.

Data Evaluation

To assess external validity of studies used in the analysis, Green and Glasgow's (2006) criteria for external validity reporting were utilized by two raters who rated all studies independently. Both raters were nurses. One had a doctorate degree while the other was completing a doctorate. Each rater read each study twice. During the first read, raters scored the studies based on Green and Glasgow's criteria (Table A2.1). To score all these studies, a simple dichotomous scale (0 = *unreported*; 1 = *reported*) similar to the TREND checklist was employed. Studies were then read for the second time to double-check initial ratings and to seek any necessary clarification.

The choice of Green and Glasgow's (2006) proposed criteria for external validity was based on recommendations by the TREND Group (personal communication, 2012). These criteria have also been previously used as a gold standard by Klesges, Dzewaltowski, and Glasgow (2008). As outlined in Table A2.1, Green and Glasgow's criteria consist of: a) reach and representativeness, b) implementation and adaptation, c) outcomes for decision making, and d) maintenance and institutionalization. Each of these four criteria comprises of several attributes which a research study must include. For the purpose of the comparative analysis, a checklist with a dichotomous rating scale was developed based on all of the 16 attributes of Green and Glasgow's four criteria for external validity reporting. The rating scale was then used to rate all TREND and non-TREND studies.

Table A2.1: External validity reported by TREND and non-TREND studies*

Criteria for External Validity Reporting**	TREND #	TREND %	Non-TREND #	Non-TREND %
I. Reach and representativeness				
Participation	7	100	6	86
Target audience	7	100	7	100
Representativeness – Settings	6	86	6	86
Representativeness – Individuals	7	100	7	100
II. Implementation and adaptation				
Consistent implementation	5	71	2	29
Staff expertise	5.5	79	4	57
Program adaptation	5	71	3	43
Mechanisms	2.5	36	0	0
III. Outcomes for decision making				
Significance	7	100	5.5	79
Adverse consequences	4.5	64	1.5	21
Moderators	5	71	0.5	7
Sensitivity	7	100	4.5	64
Costs	4	57	2.5	36
IV. Maintenance and institutionalization				
Long-term effects	2	29	2	29
Institutionalization/sustainability	5.5	79	4	57
Attrition	6.5	93	4.5	64
Mean	**5.4**	**77**	**3.8**	**54**

* Scores based on mean of two raters who independently rated all studies listed in Table A2.2.
**Criteria for external validity reporting based on recommendations of Green and Glasgow (2006).

Results and Analysis

Overall, all 14 studies included in the analysis reported non-randomized evaluations of behavioural and public health interventions. Both TREND and non-TREND studies were conducted in different parts of the world, with the majority originating from the United States. Most studies evaluated an intervention comprising of some form of education aimed at promoting healthy behaviours (e.g., smoking cessation). Target populations ranged from children to older adults, with both males and females represented. Study participants were often allocated geographically (e.g., comparing participants between two cities), while alternating allocation techniques were also employed (e.g., comparing participants within one setting but during different time periods).

All reports based on the TREND guidelines made reference to the TREND statement but lacked further discussion about its usefulness or how each of its dimensions was addressed. Of the 7 TREND studies, two studies recruited control and intervention participants in different years and one study used non-participant controls. The remaining four TREND studies were similar to all of the seven non-TREND studies in the sense that they consisted of parallel control and intervention groups which progressed simultaneously. Table A2.1 summarizes the scores and percentages of external validity reporting of both TREND and non-TREND studies based on Green and Glasgow’s (2006) criteria while Table A2.2 summarizes the extent to which TREND and non-TREND studies addressed Green and Glasgow’s criteria for external validity reporting.

Table A2.2: External validity scores of reviewed studies*

	Evaluation Score	
	#	**%**
TREND Studies		
Ciliberto et al. (2005)	12	75
Fisher et al. (2010)	11.5	72
Giangregorio et al. (2009)	13	81
Oupra et al. (2010)	12.5	78
Sorensen et al. (2010)	12.5	78
Storro et al. (2010)	13.5	84
Taylor et al. (2008)	11.5	72
Mean	**12.4**	**77**
SD	**0.75**	**-**
Non-TREND Studies		
Cardarelli et al. (2011)	9	56
Chan et al. (2011)	11	69
Elmasri (2011)	5.5	34
Kwak et al. (2009)	9.5	59
Lv & Brown (2011)	5.5	34
Ma et al. (2009)	8.5	53
Wolfers et al. (2009)	11	69
Mean	**8.6**	**54**
SD	**2.30**	**-**

* Scores based on mean of two raters who independently rated studies based on Green and Glasgow's (2006) criteria for external validity reporting outlined in Table A2.1.

Overall, all studies lacked full reporting of external validity criteria and presented limited discussion on generalizability. Across all 16 external validity criteria, mean reporting for TREND and non-TREND studies was 12.4 (SD=0.75) and 8.6 (SD=2.30) respectively. A non-parametric test (i.e., Mann-Whitney test) indicated that this difference was statistically significant (p=0.0017). To check for agreement between the scores of two raters who independently rated each study, the Intraclass Correlation Coefficient (ICC) was calculated to be 0.86 (95% CI: 0.69, 0.97). This indicated strong inter-rater reliability.

Criteria 1: Reach and Representativeness

All TREND and non-TREND studies described the target audience and compared study subjects to the intended target population while one TREND and one non-TREND study did not report on the intended settings nor compared them to those settings that declined participation. Furthermore, while all TREND studies discussed participation rates of eligible participants, one non-TREND study did not report participation rate.

Criteria 2: Implementation and Adaptation

While 5 TREND studies (71%) reported on the consistency of implementation of the various intervention components and the extent to which study settings adapted the intervention program to fit their settings, only 2 non-TREND studies (29%) reported this information. None of the non-TREND studies (0%) reported the mechanisms through which the intervention achieved its effect. This, however, was reported by a few TREND studies (36%). Moreover, most TREND (79%) and several non-TREND (57%) studies presented data on staff expertise (i.e., level of training, expertise, quality of implementation, etc.). In relation to program adaptation, 5 TREND studies (71%) reported on adaptation while 3 non-TREND studies (43%) reported this.

Criteria 3: Outcomes for Decision Making

While at least 4 TREND studies (57%) reported on all attributes of Outcomes for Decision Making, only a few non-TREND studies (21%) reported one attribute such as adverse consequences and moderator effects. Information on two attributes (i.e., sensitivity and significance) was provided by all TREND studies (100%). These two attributes were reported in several (>64%) non-TREND studies. In the TREND group, there was fair reporting of cost, moderator effects, and adverse consequences by more than half of the studies (>57%). However, these attributes were poorly reported in non-TREND studies in which less than 3 non-TREND studies reported cost (36%), adverse consequences (21%) and moderator effects (7%).

Criteria 4: Maintenance and Institutionalization

Although all studies consisted of a follow-up, only 2 TREND (29%) and 2 non-TREND (29%) studies conducted a 12-month follow-up. In TREND studies, follow-up ranged from eight weeks to two years (8-week = 1 study, 3-months = 2 studies, 6-months = 2 studies, 2-years = 2 studies) with an average of 9.7 months. In non-TREND studies, this range was from four weeks to two years (4-week = 1 study, 3-months = 1 study, 4-months = 1 study, 6-months = 2 studies, 1-year = 1 study, 2-year = 1 study) with an average of 8 months. Furthermore, several TREND (79%) and non-TREND (57%) studies reported on sustainability or evolution of the program implemented as part of the intervention. Lastly, most TREND studies (93%) and several non-TREND studies (64%) reported attrition and presented some basic discussion on reasons for why participants dropped out.

Discussion

A comparison between 7 TREND and 7 non-TREND studies is an encouraging step toward promoting external validity reporting. In this analysis, we discovered that majority of TREND and non-TREND studies did not address Green and Glasgow's (2006) criteria for external validity reporting. The TREND Group (personal communication, 2012) views these criteria as crucial for future policy decisions and knowledge translation efforts. This analysis also highlights the lack of external validity reporting in recent non-randomized intervention studies, which could limit appropriate translation of interventions to real-life situations.

In an attempt to compare external validity reporting, this analysis shows that compared to non-TREND studies (54%), TREND studies (77%) scored significantly higher on Green and Glasgow's criteria for external validity reporting. This illustrates that the use of the TREND statement promotes increased external validity reporting. Table A2.1 indicates that this difference could be due to the TREND statement's ability to draw researchers' attention toward specific external validity criteria that are important for generalizing study findings. Hence, the TREND statement, as a leap toward a systematic method of reporting, appears to promote external validity reporting in non-randomized intervention studies.

While TREND studies succeeded in reporting several criteria for external validity, there were a few areas which were not reported by several studies. These included: (a) mechanisms, (b) adverse consequences, (c) costs, and (d) long-term effects.

While the TREND statement has received criticism regarding its external validity criteria, it is important to note that the statement includes several internal validity criteria which, if reported, would also strengthen a study's external validity reporting. This is visible in Table A2.1 where non-TREND studies have performed poorly in reporting the external validity criterion of *Outcomes for Decision Making* while TREND studies have succeeded in addressing this criterion. While many attributes under *Outcomes for Decision Making* (e.g., significance, adverse consequences, and moderator effects) do not correspond to criteria under the ‘generalizability’ section of the TREND checklist, they can, however, be found elsewhere within the checklist under headings which aim to strengthen internal validity reporting.

This emphasizes that the complete use of the TREND statement can encourage the reporting of many internal validity components (e.g., significance, adverse events, implementation of intervention, moderator effects, expertise, participants, setting, cost, etc.) which can directly address Green and Glasgow’s (2006) criteria for external validity reporting. While all TREND studies made the claim that they used the TREND guidelines, they did not indicate how or to what extent. Despite this fact, improve external validity reporting among TREND studies could have resulted from the focus on several previously-discussed internal validity criteria which directly influence Green and Glasgow’s criteria for external validity reporting.

Implications

This comparative analysis between TREND and non-TREND studies has several implications for the research community. Nurse researchers considering the TREND guidelines are encouraged to thoroughly discuss how and to what extent they used the TREND guidelines, and pay particular attention to each of its criteria for external validity reporting. Nurse researchers must also realize and address these criteria when preparing study protocols before actual research is conducted. In addition to the TREND guidelines, external validity criteria of Green and Glasgow (2006) should also be considered in reports. This approach will enhance external validity reporting in journal articles and will promote subsequent knowledge translation efforts.

Dzewaltowski et al. (2004) and Steckler and McLeroy (2008) have advocated for a greater emphasis on external validity reporting among journals of applied disciplines that aim to improve the health of the public. The various characteristics of external validity recommended by these authors resemble Green and Glasgow (2006)'s external validity criteria which, along with the TREND guidelines, should be considered by researchers of all health disciplines when conducting, reporting, or evaluating non-randomized intervention studies. It is important for health researchers to report on all criteria or state that information is unavailable on criteria which may not be applicable to their research study. This can help nurses, other healthcare practitioners, and policy and administrative decision-makers to determine whether or not a given study's findings are generalizable and applicable to their local population and setting.

Findings of this comparative analysis can be used by the TREND Group to make necessary revisions to the original TREND statement to reflect external validity criteria which would emphasize and strengthen generalizability of study findings and the use of research findings in real-life situations. Since clinicians in the public health sector often conduct non-randomized research that evaluate behavioural and public health interventions, it is important for them to realize the usefulness of the TREND statement (Des Jarlais et al., 2004) and the external validity criteria proposed by Green and Glasgow (2006). Nurse researchers are encouraged to build partnerships with nurses and policy developers in order to address real-life problems and facilitate appropriate knowledge translation efforts.

Although Green and Glasgow's (2006) criteria do not focus on the type of intervention reported, the criteria do however focus on whether the treatment was consistently administered, whether there were any adverse reactions, what the cost was, what the long-term effects are, the attrition rate, how the intervention was sustained, etc. All of these are important when reporting intervention studies because they allow readers to determine whether the study findings could be generalized to their environment. Reports not addressing such criteria will make it difficult for readers to decide whether or not the study intervention can be suitable for their environment. Therefore, researchers are encouraged to utilize the TREND guidelines when reporting non-randomized intervention studies.

While the use of the TREND guidelines promotes external validity reporting, single-study results must be used with extreme caution. Should nurses and policy and administrators involved in decision-making discover that a study report is applicable to their population and setting, they must still explore and rely on synthesized results of several research studies prior to disseminating findings in the practice setting. The use of one study and its findings is insufficient and the combined results of several well-conducted studies must be considered when making decisions around the usefulness of research and its possible effectiveness in the practice setting.

Although tools to evaluate external validity reporting are useful, the final decisions around the translation of research into practice are based on judgments of healthcare professionals and policy and administrative personnel who understand the characteristics of people and settings outside the study experiment and are able to make accurate judgments about the applicability of research findings and their sustainable potential. Therefore, there is a need for creative solutions that aim to expand evidence in certain areas. As per findings of this analysis, two main areas where evidence must be expanded are in the areas of long-term follow-up of studies and intervention programs' sustainability in institutions.

Nursing and health journals that welcome reports on intervention studies are encouraged to request authors to consider the TREND statement and relevant external validity criteria before submitting non-randomized intervention study reports for publication. This can be done by incorporating this requirement into author guidelines published by the journals. Often limited funding and urgency to publish study results prevents the conduct of studies with long-term follow-up to evaluate intervention sustainability. For this reason, nursing and health journals are also encouraged to offer researchers a venue to publish follow-up reports on studies of interventions after an initial study is reported. Furthermore, funding agencies are also encouraged to consider providing increased support for long-term follow-up research studies that allow researchers to evaluate the institutionalization and sustainability of interventions (Klesges et al., 2008).

This analysis is the first attempt to compare TREND and non-TREND studies reporting non-randomized intervention studies with a control group and follow-up. Increased utilization of the TREND guidelines in the future is encouraged. This will increase the number of TREND studies which could then be used in a future analysis similar to the one presented herein. With more studies using the TREND guidelines, future reviewers will have more reports to choose from and more opportunity to set further inclusion and exclusion criteria which could ensure that selected TREND and non-TREND studies are more comparable in terms of treatment, setting and population.

Limitations

There were several limitations in this analysis. First, there were other studies that utilized the TREND statement which were not included in the analysis mainly because they employed a cross-sectional or time-series study design or were pre-post study designs without comparison or control groups. Therefore, this analysis lacks discussion on external validity reporting of these other studies that also used the TREND guidelines. Second, all TREND studies briefly mentioned, in a sentence, that the TREND statement was used as a guide. However, they failed to provide further discussion on the TREND statement which made it difficult to determine whether or not the TREND statement was useful in promoting external validity reporting of non-randomized intervention studies.

Third, although the focus of the TREND statement is not to test educational interventions, some studies used the TREND guidelines for health education interventions. While this association was not explained, there could be a future possibility where the use of the TREND guidelines could be expanded beyond health-related disciplines to include research studies from social sciences and humanities. Fourth, studies were evaluated based on a dichotomous scale while a Likert-type scale would have been more appropriate for some studies which partially met criteria for external validity reporting. Finally, the number of studies used also influenced analysis since small changes created large fluctuations in percentages. Therefore, caution is advised when generalizing findings from this analysis. With the increasing use of the TREND statement, it is recommended that a similar analysis to the one presented herein be repeated in a few years with a larger number of TREND and non-TREND studies.

Conclusion

This comparative analysis highlights the lack of external validity reporting among non-randomized intervention study reports in the medical, nursing, and public health literature. Finding from this analysis demonstrate that the use of the TREND guidelines improves external validity reporting to studies that do not use these guidelines. As a result, nurse researchers are encouraged to consider the TREND guidelines when reporting non-randomized intervention studies. It is also recommended that additional criteria for external validity reporting based on the work of Green and Glasgow (2006) be added to the TREND statement in order to promote external validity reporting by nurse researchers. Future non-randomized intervention study reports that succeed in addressing these external validity criteria will not only enhance generalizability, they will also enrich evidence-informed decision making and facilitate more appropriate translation of research findings into clinical practice.

References

Balas, E.A., & Boren, S.A. (2000). Managing clinical knowledge for health care improvement. In J. Bemmel & A.T. McCray (Eds), *Yearbook of medical informatics* (pp. 65-70). Stuggart, Germany: Schattauer Publishing.

Begg, C., Cho., M., Eastwood., S., Horton, R., Moher, D., Olkin, I., ... Stroup, D.F. (1996). Improving the quality of reporting of randomized controlled trials: The CONSORT statement. *JAMA, 276*(8), 637-639.

Caetano, R. (2004). Standards for reporting non-randomized evaluations of behavioral and public health interventions: The TREND statement. *Addiction, 99*(9), 1075-1080.

Cardarelli, K., Jackson, R., Martin, M., Linnear, K., Lopez, R., Senteio, C., et al. (2011). Community-based participatory approach to reduce breast cancer disparities in South Dallas. *Progress in Community Health Partnerships: Research, Education, and Action, 5*(4), 375-385.

Chan, W.L., Hui, E., Chan, C., Cheung, D., Wong, S., Wong, R., et al. (2011). Evaluation of chronic disease self-management programme (CDSMP) for older adults in Hong Kong. *The Journal of Nutrition, Health & Aging, 15*(3), 209-214.

Ciliberto, M.A., Sandige, H., Ndekha, M.J., Ashorn, P., Briend, A., Ciliberto, H.M., et al. (2005). Comparison of home-based therapy with ready-to-use therapeutic food with standard therapy in the treatment of malnourished Malawian children: A controlled, clinical effectiveness trial. *American Journal of Clinical Nutrition, 81*(4), 864-870.

Des Jarlais, D.C., Lyles, C., Crepaz, N., & the TREND Group. (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public health, 94*(3), 361-366.

Dzewaltowski, D.A., Estarooks, P.A., Klesges, L.M., & Glasgow, R.E. (2004). TREND: An important step, but not enough. *American Journal of Public Health, 94*(9), 1474.

Elmasri, Y.M. (2011). Effect of educational program on insight into illness and attitudes toward medications among schizophrenic patients. *Journal of American Science, 7*(3), 269-277.

Ferguson, L. (2004). External validity, generalizability, and knowledge utilization. *Journal of Nursing Scholarship, 36*(1), 16-22.

Fisher, J.R.W., Wynter, K.H., & Rowe, H.J. (2010). Innovative psycho-educational program to prevent common postpartum mental disorders in primiparous women: A before and after controlled study. *BMC Public Health, 10*, 432.

Giangregorio, L.M., Thabane, L., deBeer, J., Farrauto, L., McCartney, N., Adachi, J.D., et al. (2009). Body weight-supported treadmill training for patients with hip fracture: A feasibility study. *Archives of Physical Medicine and Rehabilitation, 90*(12), 2125-2130.

Glasgow, R.E., Green, L.W., Klesges, L.M., Abrams, D.B., Fisher, E.B., Goldstein, M.G., ... Orleans, C.T. (2006). External validity: We need to do more. *Annals of Behavioral Medicine, 31*(2), 105-108.

Glasgow, R.E., Klesges, L.M., Dzewaltowski, D.A., Bull, S.S., & Estabrooks, P. (2004). The future of health behavior change research: What is needed to improve translation of research into health promotion practice? *Annals of Behavioral Medicine, 27*(1), 3-12.

Green, L.W., & Glasgow, R.E. (2006). Evaluating the relevance, generalization, and applicability of research: Issues in external validation and translation methodology. *Evaluation & The Health Professions, 29*(1), 126-153.

Hulley, S.B., Cummings, S.R., Browner, W.S., Grady, D.G., & Newman, T.B. (2007). *Designing clinical research* (3rd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Kirkwood, B. (2004). Making public health interventions more evidence based: TREND statement for non-randomised designs will make a difference. *BMJ, 328*(7446), 996-997.

Klesges, L.M., Dzewaltowski, D.A., & Glasgow, R.E. (2008). Review of external validity reporting in childhood obesity prevention research. *American Journal of Preventative Medicine, 34*(3), 216-223.

Kwak, L., Kremers, S.P., Visscher, T.L., van Baak, M.A., & Brug, J. (2009). Behavioral and cognitive effects of a worksite-based weight gain prevention program: The NHF-NRG In-Balance project. *Journal of Occupational and Environmental Medicine, 51*(12), 1437-1446.

Lv, N., & Brown, L. (2011). Impact of a nutrition education program to increase intake of calcium-rich foods by Chinese-American women. *Journal of the American Dietetic Association, 111*(1), 143-149.

Ma, G.X., Shive, S., Tan, Y., Gao, W., Rhee, J., Park, M., et al. (2009). Community-based colorectal cancer intervention in underserved Korean Americans. *Cancer Epidemiology, 33*(5), 381-386.

Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gotzsche, P.C., Devereaux, P.J., … Altman, D.G. (2010). CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. *BMJ, 340*, c869.

Oupra, R., Griffiths, R., Pryor, J., & Mott, S. (2010). Effectiveness of supportive educative learning program on the level of strain experienced by caregivers of stroke patients inThailand. *Health and Social Care in the Community, 18*(1), 10-20.

Ross, M.W., Elford, J., Sherr, L., & Hart, G. (2004). The TREND statement: Social science, communication and HIV/AIDS. *AIDS Care, 16*(6), 667-668.

Sorensen, G., Stoddard, A., Quintiliani, L., Ebbeling, C., Nagler, E., Yang, M., et al. (2010). Tobacco use cessation and weight management among motor freight workers: Results of the gear up for healthy study. *Cancer Causes & Control, 21*(12), 2113-2122.

Steckler, A., & McLeroy, K.R. (2008). The importance of external validity. *American Journal of Public Health, 98*(1), 9-10.

Storro, O., Oien, T., Dotterud, C.K., Jenssen, J.A., & Johnsen, R. (2010). A primary health-care intervention on pre-and postnatal risk factor behavior to prevent childhood allergy. The Prevention of Allergy among Children in Trondheim (PACT) study. *BMC Public Health, 10*, 443.

Taylor, R.W., McAuley, K.A., Barbezat, W., Farmer, V.L., Williams, S.M., & Mann, J.I. (2008). Two-year follow-up of an obesity prevention initiative in children: The APPLE project. *American Journal of Clinical Nutrition, 88*(5), 1371-1377.

Treasure, E. (2004). The TREND Statement. *Evidence-Based Dentistry, 5*(4), 88-91.

Victora, C.G., Habicht, J.P., & Bryce, J. (2004). Evidence-based public health: Moving beyond randomized trials. *American Journal of Public Health, 94*(3), 400-405.

Wolfers, M.E., de Wit, J.B., Hospers, H.J., Richardus, J.H., & de Zwart, O. (2009). Effects of a short individually tailored counselling session for HIV prevention in gay and bisexual men receiving Hepatitis B vaccination. *BMC Public Health, 9*, 255.

Appendix 3: Sample Size Calculation from Pilot Study

	Pre-Test Mean (SD)	Post-Test Mean (SD)
Hypothesis Accuracy (Max score 1.0)	.131 (0.261)	.592 (0.237)
Hypothesis Expertise (Max score 1.0)	.030 (0.0548)	.142 (0.0583)
Hypothesis Breadth (Max score 3.0)	1.425 (0.712)	2.076 (0.494)

Accuracy:

- Based on expert feedback, a difference of 0.25 between control and experimental groups was suggested as educationally important (L. Martin, personal communication, 2014)
- Therefore, assuming 25% increase in hypothesis accuracy:

$$n = \frac{2\left(Z_{\alpha/2} + Z_\beta\right)^2 \sigma^2}{d^2} = \frac{2(1.96 + 0.84)^2\ 0.0622}{(0.25)^2} \approx \mathbf{16\ (per\ study\ group)}$$

Expertise: Observed difference of 0.112

- Based on expert feedback, the generation of one hypothesis was viewed as educationally important (L. Bentley-Poole, personal communication, 2014)
  - Baseline mean expertise score was 0.030
  - 1 hypothesis out of 10 possible hypothesis is 0.1
  - Therefore, 0.1 – 0.03 = 0.07 to achieve one hypothesis
- Therefore, assuming 7% increase in hypothesis expertise :

$$n = \frac{2\left(Z_{\alpha/2} + Z_\beta\right)^2 \sigma^2}{d^2} = \frac{2(1.96 + 0.84)^2\ 0.0032}{(0.07)^2} \approx \mathbf{11\ (per\ study\ group)}$$

Breadth: Observed difference of 0.651

- Based on expert feedback, a difference of 0.5 can be viewed as educationally important (T. Jewiss & L. Manankil-Rankin, personal communication, 2014)
- Therefore, assuming 0.5 difference:

$$n = \frac{2\left(Z_{\alpha/2} + Z_\beta\right)^2 \sigma^2}{d^2} = \frac{2(1.96 + 0.84)^2\ 0.37549}{(0.50)^2} \approx \mathbf{24\ (per\ study\ group)}$$

Appendix 4: Overview of Study Sessions

The structure of study sessions (in minutes) is presented below:

Study Session #1: Pre-Post Workshop (105-minutes)

5 min Welcome and introduction to study

10 min Consent & Demographics

20 min Hypothesis education

15 min Pre-Test
Generate issues and hypotheses for Scenario-1

40 min Intervention

Control	*Experimental*
Standard + Discussion	Standard + Abductive Reasoning Training

15 min Post-Test
Generate issues and hypotheses for Scenario-2

Study Session #2: Follow-up (15-minutes)

15 min Follow-Up
Generate issues and hypotheses for Scenario-3

Appendix 5: Pre-Test Study Questionnaire

Kelly, age 14

My mother caught me vomiting a few times so she insisted that I visit the nurse. It's nothing really. They were just random episodes. My mother gets over-worried. She says I don't take care of my body and that I need to eat more. I'm growing, but I don't want to be fat. Only if I didn't get so hungry all the time, I'd be slim like Heather. My boyfriend is always talking about her. My mother doesn't care about my life. She only cares about my poor grades. Sometimes, I can't sleep because I keep getting hungry. So, I just drink water. My mother calls me a zombie, which hurts my feelings. I mean… all the girls I know do the same to stay fit. I'm not the only one with these issues you know.

---

1. Identify the issues in the above care scenario:

2. Generate hypotheses which explain the situation in the above care scenario:

Appendix 6: Discussion + Standard Training Guide

The following will be covered in the 40-minute standard + discussion training:

20-Minute Discussion:

Participants will participate in an open group discussion which will be facilitated by the researcher and initiated by the following question:

"Now that you have had some learning and practice around hypotheses, let's discuss how we generate a hypothesis. What are the steps of generating a hypothesis?"

20-Minute Group Exercise:

The scenario from the Pre-Test exercise will be selected and discussed as a group. Participants will be encouraged to work as a group to come up with one hypothesis for the scenario. The researcher will facilitate this exercise while questioning participants in order to encourage the sharing of their ideas with others. The following instructions will be provided:

"Now that we have discussed the process of generating a hypothesis, let's take the scenario from the previous exercise and work together to generate one hypothesis for the scenario. This is a group task, so please share your ideas with the group."

Appendix 7: Abductive Reasoning + Standard Training Guide

The following will be covered in the 40-minute standard + abductive reasoning training:

<u>20-Minute Abductive Reasoning Training</u>

Theoretical instructions based on Vertue and Haig (2008)'s abductive reasoning process were provided to participants:

1. Phenomena detection
- Clinician collects and analyzes cues from the presenting situation

2. Inferring causal mechanisms
- Clinician identifies and groups relevant plausible causal factors and suggests their possible relationship to the detected phenomena
- Clinician uses biopsychosocial framework to structure his or her thinking process

3. Developing a causal model
- Clinician develops an illustration in which various causal mechanisms are considered

4. Evaluating the causal model
- Clinician ensures all relationships are coherent and supported by data
- Clinician conceptualizes the care situation in the most explanatorily coherent way

5. Formulating the case
- Clinician emphasizes the possible links between the various aspects of a care situation through comprehensive and integrated conceptualization of descriptive and explanatory hypotheses

<u>20-Minute Group Exercise:</u>

The care scenario from the Pre-Test exercise was used for this exercise. With some guidance, participants were encouraged to work with one another to follow the abductive reasoning steps and come up with a few hypotheses which explained the scenario. The researcher facilitated this exercise. The following instructions were provided:

"Now that I have demonstrated the steps involved in generating a hypothesis using abductive reasoning, let's work together on the care scenario from the previous exercise and apply the steps of abductive reasoning in order to generate one hypothesis for the scenario. This is a group task, so please share your ideas with the group."

Appendix 8: Immediate Post-Test Questionnaire

Tammy, age 14

I'm constantly the topic of school jokes which makes me sad. Even the friends I had talk behind my back. They say that no one likes me because I'm fat. They call me "two-eighty" because of my weight. I often sit out in gym class because my breathing gets really bad and my ankles start to hurt. In other classes, I sit at the back to avoid people. Sometimes I don't want to be at school. One girl told me that she lost weight by smoking. So recently, I started smoking but I don't want to get in trouble. My parents are really busy and barely home and I end up ordering dinner every day. When I tell them about my issues, they say that I should just tell people I'm big-boned, but that doesn't work anymore because my weight just keeps increasing.

---

1. Identify the issues in the above care scenario:

2. Generate hypotheses which explain the situation in the above care scenario:

Appendix 9: One-Week Follow-Up Test Questionnaire

Betty, age 14

Since I broke up with John last week, I've been feeling so down. I hate going to school because I can't stand my friends gossiping about us and how he dumped me because I'm starting to get fat. I've been trying to slim down but have put on another five pounds recently. I wake up feeling nauseated most mornings and feel too tired to do any exercise. Since last week, I have been secretly drinking some of my dad's scotch to numb my emotions for John and to keep my mind away from all the stress. I often feel I have no control over myself. My body has been acting weird for the last two months. Even my periods are messed up. With all that's going on in my life, it's like I'm losing track of them.

---

1. Identify the issues in the above care scenario:

2. Generate hypotheses which explain the situation in the above care scenario:

Appendix 10: Copy of Ethics Approval Letter

Hamilton Health Sciences | McMaster University | St. Joseph's Healthcare Hamilton

Hamilton Integrated Research Ethics Board (HIREB)
293 Wellington St. N., Suite 102, Hamilton, ON L8L 8E7
Telephone: 905-521-2100, Ext. 42013
Fax: 905-577-8378

August 21, 2013

PROJECT NUMBER:	**13-474**
PROJECT TITLE:	Effects of Abductive Reasoning Training on Hypothesis Generation Abilities of Level One and Level Two BScN Nursing Students
PRINCIPAL INVESTIGATOR:	**Noeman Mirza**
LOCAL PI:	**Basanti Majumdar**

This will acknowledge receipt of your letter dated August 1, 2013 which enclosed revised copies of the Information/Consent Form and the Application Form along with new Source Document submitted for the above-named study. These issues were raised by the Hamilton Integrated Research Ethics Board at their meeting held on July 16, 2013. Based on this additional information, we wish to advise your study has been given ***final*** approval from the full HIREB.

The following documents have been approved on both ethical and scientific grounds:

- The submission
- Study Protocol
- Information/Consent Form Version 2 dated August 1, 2013
- Coder Contract Version 2 dated 19/08/13
- Advertisement Poster Version 2 dated 19/08/13
- Script of Message for Web-Based System (e.g. Blackboard)
- Script for Email Message to Participant
- Script for Web-Based Message for Feasibility Study
- Source Document Only version dated May 3, 2013
- Overview of Phase-2 Study Sessions
- Hypothesis Education Module
- Picture Exercise
- Details of Discussion + Standard Training
- Details of AR + Standard Training
- Pre-Test; Post-Test and Follow-Up Test
- Demographic Data
- Experts for Content Validity of HQST and Care Scenarios
- Content Validation Packet for Experts
- Issue and Hypothesis Scoring Tool
- Picture Hypothesis Scoring Tool

The following document has been acknowledged:

- TCPS 2: Core Certificate dated 23 June 2013

Please note attached you will find the Information/Consent Form and the Advertisement Poster with Tear off Tab with the HIREB approval affixed; all consent forms/advertisement posters used in this study must be copies of the attached materials.

We are pleased to issue final approval for the above-named study for a period of 12 months from the date of the HIREB meeting on July 16, 2013. Continuation beyond that date will require further review and renewal of HIREB approval. Any changes or revisions to the original submission must be submitted on an HIREB amendment form for review and approval by the Hamilton Integrated Research Ethics Board.

PLEASE QUOTE THE ABOVE-REFERENCE PROJECT NUMBER ON
ALL FUTURE CORRESPONDENCE

Sincerely,

S. Salama

Suzette Salama PhD.,
Chair, Hamilton Integrated Research Ethics Board

HIREB #: 13-474 Page 2

Appendix 11: Copy of Study Consent Form

McMaster University

CONSENT FORM

Effects of Abductive Reasoning Training on Hypothesis Generation Abilities of Level One and Level Two BScN Nursing Students

Local Principal Investigator:
Dr. Basanti Majumdar
School of Nursing, HSC 3N28D
McMaster University
Hamilton, Ontario, Canada
905-525-9140 ext. 22456
majumdar@mcmaster.ca

Student Investigator:
Noeman Mirza
School of Nursing
McMaster University
Hamilton, Ontario, Canada
905-808-2820
mirzana@mcmaster.ca

Dear nursing student,

You are invited to take part in this study on the effects of abductive reasoning training on the quality of hypotheses generated by nursing students. Please note, this is a student research project conducted under the supervision of Dr. Basanti Majumdar. The study will help the student (Noeman Mirza) learn more about the topic area and develop skills in research design, collection and analysis of data and writing a research paper.

Why is a study on the effects of abductive reasoning training on hypothesis generation being done?

Hypothesis generation is an important part of problem-based learning (PBL). This study aims to examine the impact of an educational intervention on hypothesis generation abilities of nursing students on given care scenarios. Results of the study will help inform us about the usefulness of the intervention to undergraduate nursing education and will guide further research initiatives.

What will happen during the study?

The study consists of two parts (total of 2 hours):

Part 1: You will be asked to participate in a 105-minute study session in which you will:

1. Complete a 5-minute demographics questionnaire.
2. Participate in a 20-minute education session on what a hypothesis is.
3. Complete a 15-minute hypothesis test questionnaire (exercise #1).
4. Participate in a 40 minute group activity on hypothesis generation.
5. Complete a 15-minute hypothesis test questionnaire (exercise #2).

Part 2: You will be asked to return in a week's time and complete a 15-minute hypothesis test questionnaire (exercise #3).

Version 2: 01/08/13 **Page 1 of 5**

There will be a total of 3 hypothesis exercises. Each exercise will consist of 1 care scenario and you will be asked to generate issues and 1 hypothesis per scenario. The study will take place at McMaster University. The study will also involve a group discussion on hypothesis generation and, with your permission, handwritten notes will be taken on the content of the group discussion.

Are there any risks to doing study?

The risks involved in participating in this study are minimal. You may feel uncomfortable, anxious, or uneasy when reading the care scenarios or participating in the test-like exercises which have time limits. Since there is a group discussion component, you may also worry about how others will react to what you say.

Are there any benefits to doing this study?

Your participation in the study may improve your ability in hypothesis generation which could contribute to your performance in a PBL tutorial. We hope that what is learned as a result of this study will help us to better understand the effects of our educational intervention (i.e., abductive reasoning training) on the hypothesis generation abilities of nursing students. This could help inform future curriculum planning and guide further research.

Are there any costs or payments involved?

It does not cost you anything to be involved in this study. Student participants will receive a $10 gift certificate for Part-1 of the study and another $10 gift certificate for Part-2 of the study.

How will you keep my information private and confidential?

You are participating in this study confidentially. We will not use your name or any information that would allow you to be identified. The only persons who will know whether you participated in the study will be the researcher and the ones present at the study sessions with you. Your name and e-mail address will be noted on a separate sheet of paper for attendance purposes, and will be used to remind you about the scheduled study sessions.

On all information that will be collected, you will be asked to put the last 4-digits of your student I.D. number. In this way, all of your data will have the same 4-digit number. This will help us match your data from one study session to the next.

We will also keep all files in a locked cabinet in the Health Sciences building. Only members of the research team will have access to data during the study period. No one who reads or hears a report about this study will be able to identify you. To ensure your information is not accessible to faculty who teach you, we will make sure that faculty members from the McMaster BScN Program are not involved in the collection of data.

What if I change my mind about being in the study?

Your participation in this study is voluntary. It is your choice to be part of the study or not. If you decide to be part of the study, you can decide to stop (withdraw), at any time, even after signing the consent form or part-way through the study. If you decide to withdraw, there will be no consequences to you. Information provided up to the point where you withdraw will be kept unless you request that it be removed. In this case, you will be asked to provide the 4-digit number which you used on your questionnaire sheets and we will remove these from the study. Your decision whether or not to be part of the study will not affect your student status at the university and your continuing access to education. Choosing not to participate in this study will in no way affect your standing in the BScN program.

How do I find out what was learned in this study?

This study will enrich our knowledge about the effects of an alternative educational intervention (i.e., abductive reasoning training) on the hypothesis generation abilities of nursing students. We will share what we learn with other health care professionals and faculty through nursing and educational publications, lectures, and faculty development. If you are interested in the results of the study after publication, please contact the principal investigator and we will share results with you.

How many people will participate in the study?

We are aiming to recruit approximately 140 student participants from Level One and Level Two of the BScN program.

Effects of Abductive Reasoning Training on Hypothesis Generation Abilities of Level One and Level Two BScN Nursing Students

Local Principal Investigator:
Dr. Basanti Majumdar
School of Nursing, HSC 3N28D
McMaster University
Hamilton, Ontario, Canada
905-525-9140 ext. 22456
majumdar@mcmaster.ca

Student Investigator:
Noeman Mirza
School of Nursing
McMaster University
Hamilton, Ontario, Canada
905-808-2820
mirzana@mcmaster.ca

Please contact one of the above if you have questions or require more information about the study itself.

This study has been reviewed by the Hamilton Integrated Research Ethics Board (HIREB). The HIREB is responsible for ensuring that participants are informed of the risks associated with the research, and that participants are free to decide if participation is right for them. If you have any questions about your rights as a research participant, please call the Office of the Chair, HIREB at 905.521.2100 x 42013.

CONSENT

I agree to participate in the effects of abductive reasoning study. I have read the information presented in the information letter about a study being conducted by Dr. Basanti Majumdar and Mr. Noeman Mirza of McMaster University. I understand that my participation in this study is strictly voluntary. All information gathered for this study is strictly confidential and procedures are in place to protect my identity and my involvement. I understand that if I agree to participate in this study, I may withdraw from the study at any time. I have had the opportunity to ask questions about my involvement in this study and feel that all of my questions have been addressed. I will be given a signed copy of this form.

I would like to receive a summary of the study's results. *Yes* *No*

If yes, where would you like the results sent? Email: ________________________________

Participant's name:	**Signature:**	**Date:**
____________________	____________________	______________

Version 2: 01/08/13 **Page 4 of 5**

I observed the process of consent. The prospective participant read this form, was given the chance to ask questions, appeared to accept the answers and signed to enrol in the study.

Witness:	**Signature:**	**Date:**
____________________	____________________	______________

Investigator:	**Signature:**	**Date:**
____________________	____________________	______________

Appendix 12: Hypotheses Generated by Study Participants at Immediate Post-Test for a Care Scenario involving an Obese Female Adolescent

Control Group:

- Smoking leads to feelings of shame.
- Parental ignorance can lead to poor life choices.
- Obesity increases your risk of inactivity.
- Poor family support can influence a child’s decision / physical development.
- Following certain lifestyle choices will lead to poor health long term.
- Peer pressure can cause depression.
- For school aged children smoking is associated with losing weight.

Experimental Group:

- Obesity can cause low self esteem in middle aged school children.
- Individuals who are bullied and are excluded from groups have higher chances of becoming depressed and/or overweight.
- Unhealthy eating and lack of physical activity will induce childhood obesity.
- Bullying and name calling can lead to social isolation.
- Smoking causes respiratory problems.
- Obesity can cause poor body image in adolescents.

Appendix 13: Hypotheses Generated by Study Participants at One-Week Follow-Up for a Care Scenario involving a Teen Pregnancy

Control Group:

- Sickness is caused by stress and lack of control.
- Emotions can easily overwhelm adolescents and lead to avoidable consequences.
- Adolescent girls who experience high stress are at increased risk of having irregular menstruations.
- Stress in a female teenager causes abnormal body responses, such as abnormal periods and feelings of malaise.
- Sex can lead to pregnancy.
- Teen relationships may lead to pregnancy.
- Relationship issues are likely to be associated with weight gain, fatigue and nausea.
- Girls who break up with their boyfriends gain weight and become depressed.

Experimental Group:

- A client that is pregnant will experience an increased degree of emotional stress.
- When individuals are put in stressful situations such as break ups, they may turn to unhealthy coping mechanisms such as drinking.
- Pregnancy can lead to fatigue, weight gain, and nausea.
- Pregnancy can lead to menstrual changes.
- Due to Betty’s negative coping strategies she is harming her body and the infant by depending on alcohol to cope with her break up.
- Missing your period, gaining weight and experiencing morning sickness can be indicative of being pregnant.