A COMPARISON OF CARDIOVASCULAR NURSES’ SIMILARITY JUDGMENTS

A COMPARISON OF CARDIOVASCULAR NURSES' SIMILARITY JUDGMENTS AT TWO POINTS IN THE NOVICE-EXPERT CONTINUUM

By SALLY BINKS, BScN

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment

of the Requirements for the Degree of Master of Science

McMaster University MASTER OF SCIENCE (2011) Hamilton, Ontario (Nursing)

TITLE: A Comparison of Cardiovascular Nurses’ Similarity Judgments at Two Points in the Novice-Expert Continuum AUTHOR: Sally Binks, BScN SUPERVISOR: Dr. Ruta Valaitis

NUMBER OF PAGES: x, 139

Abstract

Safe nursing care of patients with cardiovascular disease requires specialized knowledge and skill. However, education for such specialty practice has little empirical basis. Research has suggested that novices perceive primarily surface aspects of a domain[1] while experts perceive primarily its deep principles. Insight into such perceptual differences may help inform future pedagogical strategies. The purpose of this study was to investigate whether or not such differences are found in the domain of cardiovascular nursing.

A forced choice triad task was developed to test the study hypothesis. Each triad trial consisted of a target item and two referent items. Referent items were designed to be similar to target items based on surface (S) features such as patient age or sex, deep (D) features such as the goal of therapy, or both surface and deep features (S/SD). Participants were asked to choose the referent item that they believed was more similar to the target item and to provide a rationale for their choice. The instrument was completed by n=17 nurses in two practice cohorts: a novice/less experienced cohort with less than three years of experience in cardiovascular nursing and an experienced/competent cohort of Nurse Practitioners with at least ten years of experience in the specialty.

Statistically non-significant differences in scores between study cohorts occurred in the hypothesized direction. However, due to the small final sample size, no definitive conclusions could be made based on this study's findings.

[1] The term domain, as used in this document, refers to "a sphere of knowledge, influence, or activity" (Merriam-Webster, 2011).

Acknowledgments

I would like to thank the members of my supervisory committee for persevering with me through this long journey. Dr. Ruta Valaitis has made available to me so many opportunities for learning and growth, I cannot count them all; I thank her for these, as well as for her warm and generous spirit and her endless enthusiasm. I would like to thank Dr. Patricia Strachan for the contribution of her expertise in cardiovascular nursing, and for the insightful questions she raised in the review of the many drafts of this document. Dr. Ilana Bayer's sharp eye for typographical errors and misplaced punctuation was as appreciated as her encouraging manner and her subject matter expertise. I cannot thank Dr. Kevin Eva enough for the contribution of his expertise in any number of areas and for his commitment to continue as a committee member well after his departure from McMaster University. Thanks to Dr. Sandra Carroll for her contribution as the external examiner for my defence.

I would like to thank the content matter experts from across Canada who assisted in the review of the study instrument: Mr. Rodolfo Pike, Ms. Annemarie Kaan, Dr. Martha Mackay, and Ms. Sandra Lauck.

Many thanks to Dr. Karen Harkness for her assistance as an independent rater of participant responses, and for her interest in and support of my work generally.

The McMaster University Nursing Graduate Program is due thanks for its financial support in the form of an internal scholarship. Former Assistant Dean of the Nursing Graduate Program Dr. Margaret Black was a quiet but effective supporter who was always willing to lend a sympathetic ear and practical assistance. It was always a pleasure to walk into 3H48 to see Ms.

Doris Hutchison at her desk; her genuine joy in seeing students succeed and her many kindnesses will never be forgotten.

Thanks to Ms. Sarah Devantier of the Categorization Lab at the University of Western Ontario for her always prompt, cheerful, and helpful responses to queries about the forced choice triad task method. I will be sure to cite you “lots”!

Mr. Oliver Klimek of the McMaster University Student Research Ethics Committee went above and beyond the call of duty with his advice and assistance.

Finally, thanks to my good friends Mr. Don Wildfong and Ms. Lina D’Onofrio for their unflagging encouragement and support throughout my graduate studies.

Table of Contents

Figures and Tables

Chapter One

Introduction

Hospital inpatient acuity and complexity are high and informed nursing surveillance is necessary for the early recognition of complications and changes in patient status. Therefore, the educational preparation for specialty nursing practice, such as cardiovascular nursing, should be rigorous and based on proven efficacy. An understanding of differences in perception at increments along the novice-expert continuum in nursing may help inform pedagogical strategies for specialty education. Research in a variety of domains has suggested that perception in complex subject domains differs across the novice-expert continuum, with novices focusing on superficial or surface aspects of the domain and experts focusing on the domain's deep, structural or relational principles[2]. The purpose of this study was to investigate whether or not such differences are found in the domain of cardiovascular nursing.

In Chapter One, rationales for conducting research to inform specialty practice in cardiovascular nursing and for investigating differences along the novice-expert continuum are offered. A summary of critiques in the nursing literature of Patricia Benner's highly influential model of novice-to-expert nursing practice is provided. Methodological issues in knowledge elicitation are explored. Finally, the use of similarity

[2] The term domain, as used in this document, refers to "a sphere of knowledge, influence, or activity" (Merriam-Webster, 2011).

judgments and of the forced choice triad task, a method that entails similarity judgments, to investigate novice-expert differences is discussed.

In Chapter Two, the research purpose and questions are identified and the phases of the development of the forced choice triad task used in the study are described. The findings of the study, although not statistically significant, showed differences in the hypothesized direction and are reported in Chapter Two.

In Chapter Three, the provisional implications of the study findings for future research in specialty nursing education are discussed and the barriers to participant recruitment experienced during the study are described; recommendations to researchers and organizations to mitigate these barriers are made.

A Rationale for Conducting Research to Inform Cardiovascular Nursing Education

The prevalence of cardiovascular disease (CVD) in Canada remains high, ranging from 15% to 30% in the over-sixty age group (Manuel, Leung, Nyugen, Tanuseputro, & Johansen, 2006). Between 1969 and 1997, there was a 50% decline in the death rate for CVD which is largely attributable to more effective treatment modalities (Manuel et al., 2006). Increasingly, patients are managed in the community and procedures that would have required hospital admission in the past, such as angioplasty, are now being performed on an outpatient or short-stay basis (Health Canada, 2006). Patients who present with acute episodic conditions or with an acute exacerbation of a chronic condition are likely to be admitted to hospital. Nurses who work in acute care facilities, therefore, are more likely to care for the most unstable and complex CVD sufferers.

Clarke and Aiken (2003) identify two chief nursing functions in the care of hospitalized patients: surveillance and taking appropriate action when significant changes in status are observed. Kutney-Lee, Lake and Aiken (2009) describe nursing surveillance as consisting of continuous observation, recognition of significant cues, interpretation of information, and decision-making. Clearly, such activities are informed by specialized knowledge. The Canadian Council of Cardiovascular Nurses (2008) enunciates in its standards the competencies required for practice in the specialty, including the use of knowledge of cardiovascular anatomy, physiology and pathophysiology to make decisions and the use of specialized skills to perform assessments and provide interventions.

Nurses enter practice as generalists (College of Nurses of Ontario, 2008); education beyond that provided in entry-to-practice programs is required to prepare them to practice in clinical specialities such as cardiovascular nursing. This specialty education is generally provided in the clinical setting during orientation programs and through continuing education programs offered by community colleges. Specialty nursing education; however, has little empirical foundation, owing in part to the paucity of research on the topic. The literature reporting on staff development and orientation is dominated by descriptive accounts (Chesnutt & Everhart, 2007; Collins & Thomas, 2005; Everson, 2005; Gavlak, 2007; Graham, 2006; Harrelson, 2007; Lavoie-Tremblay et al., 2002; Lott, 2006; Marcum & West, 2004; McKane, 2004; Nolan & Murphy, 2006; Penphrase, 2000; Peterson & Van Buren, 2006; Proulx & Bourcier, 2008; Ragsdale & Mueller, 2005; Rauen, 2004; Santucci, 2004; Wolf, 2005; Wong, 2006; Zekonis & Gantt, 2007). Publications on specialty education that include an evaluative component are frequently based on self-reports (Billings, Jeffries, Daniels, Rowles, Stone & Stephenson, 2006; Wray, Whitehead, Setter & Treas, 2006) or on questionable evaluative methods (Wilgis & McConnell, 2008). In the few reports in the nursing literature of evaluations based on objective criteria, the outcomes of interest tend to be retention rates or cost savings rather than educational efficacy, quality of practice, or nursing-sensitive patient outcomes (Cavanaugh & Huse, 2004; Loiseau, Kitchen & Edgar, 2003; Pine & Tart, 2007; Squires, 2002).

An understanding of the process whereby expert practice in a nursing specialty develops may inform the pedagogical strategies used in specialty nursing education. Differences in performance in a domain reside not only in the extent of domain knowledge but in the way that knowledge is structured and stored in memory (Lajoie, 2003; Patel, Arocha & Kaufman, 2001; Regehr & Norman, 1996). Over time and with deliberate effort, individuals are thought to develop highly elaborated and well-organized knowledge structures that enable them to efficiently retrieve the appropriate information from long-term memory when cued by some stimulus (Gentner, Rattermann, & Forbus, 1993; Patel, et al., 2001; Schmidt, Norman, & Boshuizen, 1990). By reducing cognitive load, it is posited that this efficient retrieval enables the domain expert to quickly identify and attend to relevant information and to focus mental resources upon problem-solving (Horgan, Millis, & Neimeyer, 1989; Patel et al., 2001; Schmidt, et al., 1990). Holyoak and Koh (1987) suggest that both novices and experts use a similar mechanism of retrieval to solve novel problems, but that experts are more adept than novices at perceiving those features of the novel situation that are relevant to the attainment of a goal and using them as retrieval cues.

A variety of models seek to describe the representation of domain knowledge (Custers, Regehr, & Norman, 1996); all of these models assume that the retrieval of knowledge from memory entails a process of judging similarities; i.e., of recognizing and comparing relevant features of a current scenario to those encoded in memory. This study sought to investigate whether or not differences in the basis of similarity judgments are observed at increments along the novice-expert continuum in the domain of cardiovascular nursing.

A Rationale for the Study of Differences along the Novice-Expert Continuum

The rationale for studying differences at increments along the novice-expert continuum in perception, memory and performance is based on a relative account of expertise. In the relative view, expertise is attainable by almost anyone who is willing to invest the time and effort necessary to master the knowledge and skills of a domain (Chi, 2006). If the relative account is valid, it follows that strategies to enhance learning and performance could be developed based on an understanding of differences observed between individuals at increments along the continuum. It also follows that such strategies should acknowledge the particular knowledge and skills of the domain (Corcoran-Perry, Narayan, and Cochrane, 1999; Norman, 2006).

Lajoie (2003) advocates studying novice-expert differences in order to 1) delineate the trajectory from novice to expert; 2) develop learning assessment methods; and 3) support learning in a way that is appropriate to the learner's position on the novice-expert continuum. Novice-expert differences in learning and in the performance of skills--whether they be psychomotor or mental--suggest that different educational strategies should be used in both areas. Once the automatic level of performance, characterized by minimal conscious effort, is reached the individual is less able to monitor his or her performance and to make deliberate adjustments to it (Ericsson, 2006a). The role of the mentor or teacher at this stage, then, is to devise learning tasks that are outside the learner's current competence and to provide feedback on performance (Ericsson, 2006a). Lajoie (2003) asserts that progress toward expertise can be accelerated by rendering that trajectory visible to learners through explicit exemplars of expert performance and by providing feedback that compares the learner's performance to expert performance on a task. Patel, Glaser, and Arocha (2000) argue that learning is increasingly under the control of the learner as he or she advances and suggest that an important aspect of professional education is the teaching of self-regulation or metacognition. In a study investigating the learning processes of novice and expert nurses, Daley (1999) found that expert nurses were more aware than the novices of their learning processes; she also suggested that the development of expertise in a domain might require the development of metacognitive strategies.

Facility in analogical problem-solving has been attributed to the ability to recognize similarities between the deep principles underlying problems in different contexts; whereas, poor ability to transfer analogies between contexts is attributed to similarity judgments being limited to surface features of problems (Chi, Feltovich, & Glaser,1981; Quilici & Mayer, 1996; Regehr & Norman,1996). Teaching that explicitly calls novices' attention to the deep principles underlying a domain has been shown to help improve learners' understanding of important domain concepts and improve analogical transfer (Eva, Neville, & Norman, 1998; Goldstone, 2006; Hardiman, Dufresne, & Mestre, 1989).

Models of the Novice-Expert Continuum

Dreyfus and Dreyfus (1984) proposed five stages in the trajectory from novice to expert performance in a domain: novice, advanced-beginner, competent, proficient, and expert. Novices, according to the Dreyfus model (1984) rely on context-free rules based on facts and features to guide their actions. For example, novice nurses would be more likely than nurses at more advanced levels to assess the adequacy of blood pressure based on its numerical value rather than on the basis of whether the particular blood pressure observed is providing adequate end-organ perfusion. Dreyfus (2004) ascribes novice performance to this context-free application of rules. The advanced-beginner's practice is also governed by rules; however, the rules are applied with some recognition of their relevance in a particular instance. This recognition arises from exposure to a number of instances, enabling the advanced-beginner to recognize similarities between the same feature set in different contexts (Dreyfus, 2004). The advanced-beginner, while able to recognize these similarities, often cannot judge the significance of them in relation to the situation as a whole (Dreyfus & Dreyfus, 1984). For example, an advanced-beginner might recognize fever in post-operative patients but might not realize that fever at different points in the post-operative course may occur for different reasons, some of which are amenable to independent nursing actions and some of which require the intervention of other disciplines. The competent practitioner, in contrast, demonstrates a beginning ability to judge the salience of situational features and to prioritize effectively through deliberate planning (Dreyfus, 2004). At the competent level of performance, the role of emotion becomes important: Dreyfus (2004) suggests that practitioners' emotional responses to the consequences of their decisions provide the impetus to continue to develop. Competent practitioners predominantly use an analytical, versus, heuristic, approach to problem-solving, whereas proficient practitioners' decision-making is guided by a holistic perception of a given situation and its nuances (Dreyfus & Dreyfus, 1984). Like the proficient practitioner, the expert perceives situations holistically; expert practice is distinguished from proficient practice by its speed and fluidity, its largely unconscious nature and by the expert's ability to make subtler discriminations (Dreyfus & Dreyfus, 1984; Dreyfus, 2004).

Patricia Benner adapted the Dreyfus Model to explicate the novice-expert continuum in nursing practice in her seminal study (Benner, 2001). Using Critical Incident Technique, in which participants were asked to describe “An incident in which you feel your interventions really made a difference in patient outcome... [and] that you think captures the quintessence of nursing” (Benner, 2001, p. 300) Benner conducted individual and small group interviews with nurses at different levels of experience in a variety of acute care hospital settings. Although Benner’s description of the characteristics of each of the five practice levels closely resembles those of the Dreyfus model, the nursing literature tends to ascribe the theory or model to Benner. The model has been extremely influential, having been widely used as a framework in undergraduate nursing education programs (Carlson, Crawford, & Contrades, 1989; Gatley, 1992; Larew, Lessans, Spunt, Foster & Covington, 2006; Meyer & Xu, 2005; Waldner & Olson, 2007), staff development (Gatley, 1992; Henning & Cohen, 2008; Shirley, 2007; Storey, Crist, Nelis, Murphy, & Fisher, 2007; Square, 2010) and performance evaluation (Bourgault & Smith, 2004; Cote & Burwell, 2007; Shapiro, 1998; Tolhurst & Bonner, 2000).

Despite its wide influence, significant criticisms have been made of Benner’s work. Gobet and Chassy (2008) consider Benner’s use of the interview method inappropriate, given her claim that much of experts’ knowledge is tacit and ineffable. Researcher bias is evident in Benner’s stated purpose for conducting the study:

> Lack of public recognition of the nature of nurses' knowledge causes nurses' clinical learning to be neglected in local practice settings. This work was the basis of a movement to make local experiential learning public and accessible by reflecting on the clinical knowledge that is evident in narratives of experiential learning (Benner, 2001, vii).

Thus, this research, rather than being purely investigative in nature, appears to have had the objective of making nursing knowledge and work visible. This purpose has been lauded by some authors; Rudge (1992) remarks, "...one could argue that accentuating the positive aspects of care acts as a political affirmation of nursing's worth" (Rudge, 1992, p. 85). Allen and Cloyes (2005) remark that such experiential research is often used to study those who are seen as "marginalized" in order to give them voice. Cash (1995) points out that a significant number of the practice exemplars reported by Benner concern disputes between physicians and nurses that hint at medicine's hegemony over nursing: "The only point in publishing [a description of a critical incident] is if it is out of the ordinary. In this case, the nurse was the victor in the power struggle, therefore it is worthy of publication" (Cash, 1995, p. 533). A second, unstated goal of this research, therefore, appears to have been to elevate nursing's status relative to that of medicine.

Significant bias is also evident in Benner's use of the Critical Incident Technique in that participants were given explicit instructions as to the nature of the clinical experiences that were to be discussed in interviews. As Altmann (2007) states, Benner demonstrates a positive bias in that the published narratives emphasize exemplary practice that consistently results in successful outcomes. Nelson and McGillion (2004) condemn the use of nursing narratives as a research method on the same grounds:

> We make the case that the chosen solution to the problem of visibility in nursing practice, narrative production, has led to particular constructions of practice...Rather than viewing these narratives as processes that illuminate the world of practice, we argue that they produce a palatable and highly desirable discourse about practice for a nursing audience (Nelson & McGillion, 2004, p. 637).

In reference to Benner's study in particular, Nelson and McGillion (2004) assert that Benner's subjects were coached to "[edit] their thoughts and concerns in line with the interviewer's narrative agenda" (p. 637). Such an approach to research confounds the aim of inquiry, which is to discover the truth about some phenomenon.

Benner's participants were identified on the basis of supervisor and peer nomination (Benner, 2001). Practice level was not assigned a priori to individuals; instead, the clinical practice exemplar was classified as reflecting practice level. Cash (1995) points out an apparent incongruency between Benner's methods and her conclusions that she fails to reconcile: Although Benner claimed that the unit of analysis was not the individual but the practice exemplar, she posited a developmental model of skill acquisition that implies incremental changes in individual performance over time. Benner has been criticized by some for failing to provide an operational definition of expertise (English, 1993; Ericsson, White, & Ward, 2007) and by others for failing to address the social, cultural and political issues that bear on the identification of experts (Cash, 1995; Paley, 1996). Both Paley (1996) and Cash (1996) argue that the identification of expertise entails the consensus of relevant parties and that the issue that must be addressed is that of who the relevant parties are. Benner's failure to describe the specific mechanisms by which an individual advances along the novice-expert continuum has also been criticized (English, 1993).

English (1993) criticizes Benner's and Dreyfus' failure to provide a clear account of intuition in expert practice; he disputes their claims that intuition is a holistic or gestalt phenomenon by asserting that it is, in fact, characterized by features accounted for in models of cognition. Ericsson, Whyte, and Ward also (2007) dispute the use of intuition by experts in problem-solving; instead, these authors claim, experts form mental representations of information in working memory in order to consider, refine, and generate solutions. Chunking, or the grouping of units of behaviour or information, is the mechanism proposed by Rolfe (1997) to account for ostensibly intuitive expert performance. Cash (1995) and Paley (1996) challenge the notion that intuition, if it exists, is used only by experts. Chi (2006) and Paley (1996) further suggest that, because it incorporates deeply ingrained habits of thought and perception, intuition may actually hinder performance on occasion.

Dall'Alba and Sandberg (2006) criticize stage models of professional development such as those of Dreyfus and Benner for their emphasis on attributes that characterize practice at different levels in a decontextualized manner; they contend that the professional's understanding of what constitutes his or her practice in vivo determines how practice develops over time. These authors argue that both studies were cross-sectional rather than longitudinal and that, therefore, the assumption of step-wise skill acquisition that they made was not justified (Dall'Alba & Sandberg, 2006).

An assumption underlying the present study is that a novice-expert continuum of nursing practice exists. Two points within that continuum, defined on the basis of criteria derived from the expertise literature, are examined using a methodology that is less vulnerable to researcher bias than that employed by Benner.

Novice-Expert Differences in Memory and Perception

Several frameworks subsuming different theories of clinical knowledge representation have been advanced. In theories that fall under a prototype framework, a particular disease or clinical situation is believed to be represented in memory as a set of typical features or as a prototypical exemplar (Custers et al., 1996). Under the instance-based framework, conversely, individual clinical situations are retained intact in memory (Custer, et al., 1996). Under a semantic network framework, the structure of knowledge in memory is seen as a network of "nodes" of concepts that are linked in a meaningful way; knowledge is retrieved via pathways through the links between nodes that are perceived as relevant in a given situation (Custers, et al., 1996). Whether one accepts a prototype model, an instance-based model, or a semantic network model of clinical knowledge representation (Custers, et al., 1996), the process of judging similarities between aspects of a current scenario and information encoded in memory has a central role in the efficient retrieval of relevant domain knowledge (Feldon, 2007; Horgan et al., 1989). Research in a variety of domains has suggested that the perception of problem spaces in complex subject domains differs across the novice-expert continuum, with novices primarily perceiving superficial, surface, or obvious aspects of a domain and experts primarily perceiving the deep, relational or structural domain principles in the generation of problem solutions (Chi et al., 1981; Devantier, Minda, Hadarra & Goldszmidt, 2008; Hmelo-Silver, Marathe & Liu, 2007; Hmelo-Silver & Pfeffer, 2004; Rabinowitz & Hogan, 2002).

Methodological Issues

Classification of Practice Level

The initial challenge in the study of differences at increments along the novice-expert continuum is the classification of practice level. Arguably, this difficulty is amplified as the expert end of the continuum is approached; a novice to a domain is relatively simple to identify and define, whereas the identification of domain experts is a matter of contention. In their critique of Benner’s work (2001), Gobet and Chassy (2008) argue that the criteria Benner used to classify practice level (practice longevity and supervisor nomination) are invalid because they do not consistently correlate with performance on domain tasks. Ericsson (2006a) challenges the validity of reputation, peer nomination and practice longevity for the same reasons. He proposed that experts be identified on the basis of reproducible, superior performance on domain-specific tasks

(Ericsson, 2006a). Farrington-Darby and Wilson (2006) offer a possible explanation for the invalidity of peer nomination and position as criteria for expert selection when they suggest that people may be identified as experts for reasons that are unrelated to expertise. Daley (1999) also impugned longevity and nomination as practice level criteria but in her study she classified nurses' practice level on the basis of the conformity of the themes in their practice narratives to Benner's model; a strategy that appears to endorse Benner's findings while condemning her methods. Despite the concerns about the validity of these criteria, many nursing authors have used practice longevity and peer, supervisor or even self-nomination as criteria for classifying practice level (Currey & Botti, 2006; Ebright, Patterson, Chalko, & Render, 2003; Lyneham, Parkinson, & Denholm, 2008; Reischman & Yarandi, 2002; Tabak, Bar-Tal, & Cohen-Mansfield, 1996; Taylor, 2002). Other investigators have used these criteria in combination with professional designations or certifications to identify experts (Aitken, 2000), but no nursing studies have used Ericsson's suggested approach. Farrington-Darby and Wilson (2006) advocate specialization within a domain as one criterion for selecting experts, particularly when the specificity of domain knowledge is of interest. Henderson-Everhardus (2004), in an analysis of differences in the accuracy of detection of vascular insufficiency among a group of 76 nurses, found that only those nurses with specialty certification in vascular nursing demonstrated reliably superior performance in identifying vascular insufficiency through physical assessment; the findings of which were validated by diagnostic imaging. Ung et al. (2002) found that specialty certification was positively correlated with superior performance in intravenous cannulation by nurses; this finding was replicated in a study by Jacobson and Winslow (2005). The use of professional designations and domain specialization to identify experts a priori, however, can be criticized on the grounds that these are evidence only of the attainment of a standard of performance at a level more advanced than that of the beginning or intermediate practitioner, rather than of expertise.

Gobet and Chassy (2008) raise a question with important implications for studies of novice-expert differences: Do individuals demonstrate practice at different levels for different tasks within a domain? If so, perhaps performance in specific domain tasks rather than individuals should be the subject of study of novice-expert differences. If such an approach is taken, the identification of tasks that are representative of the domain in question and the criteria defining performance on these tasks become the challenge, rather than classification of practice level.

In the present study, a novice/less experienced cohort was defined based on practice longevity of less than three years in cardiovascular nursing, as practice longevity is assumed to be a necessary but insufficient condition for development of expertise in a domain. An experienced competent cohort was defined on the basis of practice longevity in cardiovascular nursing of at least ten years, advanced educational attainment, and satisfactory performance on objective structured clinical examinations. The rationale for the use of these definitions in this study is that the experienced competent cohort could reasonably be expected to perform more like experts than novices on a series of similarity judgments.

Knowledge Elicitation

The ideal knowledge elicitation methodology has been described as one that is objective and reliable, does not bias responses, and produces a complete and accurate representation of the subject's knowledge structure (McClure, Sonak, & Suen, 1999). It has been suggested that more than one method should be used to produce a complete account of an individual's domain knowledge, since methods may differ as to the kinds of knowledge that they elicit; i.e., procedural/declarative knowledge or tacit/explicit knowledge (Hoffman, Shadbolt, Burton, & Klein, 1995). However, the findings of research on this "differential access hypothesis" have been inconsistent (Hoffman & Lintern, 2006).

Cook (1994) classifies methods for knowledge elicitation into three categories: observations and interviews, process tracing, and conceptual techniques. The advantages of observation include the minimization of disruption to care and of discrepancies between what participants report and what they actually do (Pope & Mays, 1995). A potential hazard of this method is the Hawthorne effect or the tendency for individuals who are being observed to perform differently than they would if unobserved (Pope and Mays, 1995). Another limitation to observation is that the observer is necessarily selective as to the data recorded (Currey & Botti, 2006; Pope & Mays, 1995).

The interview enables the collection of data that are not directly observable (Patton, 2002). A limitation to the retrospective interview technique is that recall may be incomplete or subject to conscious or unconscious post hoc re-interpretation (Lincoln &

Guba, 1985). Allen and Cloyes (2005) assert that language is a social practice and that people speak differently under different circumstances. When language is treated as a neutral medium, the validity of what is said is unquestioned and the context within which the interview is conducted is seldom addressed (Allen & Cloyes, 2005). Interviews, these authors claim, are a “complex social performance” (Allen & Cloyes, 2005, p.100) and, therefore, the production of language--how people talk about their experiences--should be the object of analysis in research using interview data.

The process tracing category includes concurrent and retrospective verbal reports on task performance, protocol analysis and decision analysis (Cook, 1994). In concurrent verbal reporting or “think-aloud” protocols, people verbalize their thought processes as they carry out a task. A major advantage of concurrent reports over retrospective reports is that there is no delay between processing and reporting in concurrent reports (Yoshida, 2008). The accuracy of recall may vary inversely with the interval between task performance and report, while longer thought sequences are vulnerable to inferential bias since they are difficult to recall accurately and completely (Ericsson, 2006b). Furthermore, the findings of concurrent and retrospective reports on the same task appear to differ (Eva, Brooks & Norman, 2002). Reconstructed methods of problem solving; that is, the method that the individual reports having used to solve a problem, may differ from the method that was actually used (Woolery, 1990). The veridicality of think-aloud protocols have been challenged for a number of reasons. *Veridicality* refers to the fidelity with which verbal reports reflect cognitive processes (Kirk & Ashcraft, 2001). Poor verbal facility is an obvious threat to veridicality (Yoshida, 2008). In the clinical setting, the presence of conscious patients might constrain verbal reporting. Of particular relevance to the study of expert cognition is the notion that cognitive processes that have become automated cannot be verbalized (Kirk & Ashcraft, 2001). This idea is challenged by Hoffman and Lintern (2006), who assert that experts can, when prompted, explicate their thinking. In Aitken's study of critical care nurses' clinical data processing (2000) seven of eight participants were able to describe the salience of their observations to their assessments and interventions. Likewise, Crandall and Getchell-Reiter (1993), using interview probes, elicited specific assessment cues used by neonatal intensive care nurses in critical patient care situations that were not spontaneously articulated in initial interviews. The veridicality of the think-aloud method has also been questioned on the grounds that it forces the verbal expression of non-verbal processes and because people are frequently unconscious of the influences upon their thought processes (Eva et al., 2002; Medin, 1989).

The possibility that concurrent verbal reports induce reactivity is another criticism that has been made of this method (Kirk & Ashcraft, 2001). *Reactivity* refers to the risk that the act of verbal reporting may change the cognitive processes described (Kirk & Ashcraft, 2001). Ericsson and Simon (1980) attribute reactivity in interview data to errors on the part of interviewers in eliciting verbal reports; they proposed a protocol analysis technique for interviewing in which subjects verbalize their thoughts as they complete a well-defined task without interpretation, analysis or explanation of their thoughts in order to avoid reactivity effects.

Conceptual techniques seek to elicit mental representations of domain concepts and their relationships to one another, and include methods such as concept mapping (Cook, 1994). Concept mapping is a technique for producing a concrete representation of knowledge structure. In a concept map, domain concepts are represented by “nodes”; the strength and direction of relationships between concepts are represented by lines between nodes called propositions (Ruiz-Primo & Shavelson, 1996). Concept maps can be created by the individual whose knowledge representation is under study, or by an investigator on the basis of interview or written data (Ruiz-Primo & Shavelson, 1996). Domain concepts may be supplied by the investigator for the subject to organize, or they may be generated by the subject. Evaluation of concept maps entails either some scheme for scoring the number of nodes/propositions and the accuracy of the described relationships between nodes or comparison to a criterion map that is assumed to reflect a domain expert’s knowledge structure (Ruiz-Primo & Shavelson, 1996). The poor correlation between scores on standardized tests of domain knowledge and scores on concept maps found by West, Pomeroy, Park, Gerstenberger, and Sandoval (2000) and the lack of consistency among concept maps of experts within a domain found by McGaghie, McCrimmon, Mitchell, and Thompson (2004) challenge the concurrent and construct validity of concept maps in the elaboration of knowledge structure. Schmidt (2004) suggests that concept maps provide a representation of the use of knowledge in specific instances rather than a stable representation of knowledge structure.

Similarity Judgments as a Method of Investigating Differences along the Novice-Expert Continuum

Several knowledge elicitation methods (forced choice triad tasks, repertory grid technique and multidimensional scaling) require participants to make similarity judgments. The construct of similarity has been invoked in accounts of concept formation, analogical reasoning, inference generation, prediction, memory and problem-solving (Genter & Markman, 1994; Goldstone & Son, 2005; Sloman & Rips, 1998; Taylor & Hummel, 2009; Tversky, 1977). However, it has been argued that similarity may be too flexible to account for any cognitive process, since judgments of similarity vary according to the context in which a comparison is being made, the objects in a stimulus set, and the direction of the comparison (Goldstone & Son, 2005; Medin, Goldstone, & Gentner, 1993). Tversky's diagnosticity principle maintains that the grouping of objects in a stimulus set influences similarity judgments (Tversky, 1977). Medin et al. (1993) express a similar notion: They assert that the context within which a comparison is made may suggest the salient properties for comparison; the degree to which entities share these properties increases the similarity between them. The same authors also argue that similarity judgments are directional: The statement "a surgeon is similar to a butcher" has a different connotation than "a butcher is similar to a surgeon" (Medin, et al., 1993, p. 259). Tversky (1977) asserted that similarity judgments are asymmetrical, since similarity statements consist of a more salient subject and a less salient referent. Gentner (1988) describes a “relational” shift that occurs developmentally: Children tend to make similarity judgments on the basis of superficial or attributional properties whereas adults tend to make similarity judgments based on “relational” or functional properties (Genter, 1988). This shift is ascribed to the acquisition of domain-specific knowledge, since it has been demonstrated to occur earlier in domains that are familiar to children (Gentner, 1988). Thus as Medin et al. (1993) argue, although similarity is flexible, similarity judgments can be shown to vary in a systematic way. Medin remarks, “Although similarity must function at some level in the induction of concepts, the induced categories are not ‘held together’ by the undifferentiated force of similarity, but rather by structural principles”(Medin, 1989, p.1476).

Falan, in a concept analysis of similarity, defined similarity as “the correspondence between objects, events, and/or circumstances based on features, dimensions, and relationships” (Falan, 2010, p. 147). She identified four levels of similarity: surface, literal, thematic, and analogy (Falan, 2010). Surface similarity entails a similarity between objects on the basis of superficial features such as colour (Falan, 2010). In literal similarity there is correspondence between superficial features as well as in the relationships between the features of entities (Falan, 2007). Thematic similarity entails a relationship between two superficially dissimilar entities; e.g., the similarity between “food” and “life” is that food sustains life (Falan, 2010, p. 147). Analogy consists of an inferred similarity between relationships connecting features in different domains (Falan, 2007). Holyoak and Koh (1987) describe *surface* and *structural* similarity; the distinction lies in whether or not features of two entities influence goal attainment in problem states. Attributional similarity, as described by Medin et al. (1993) is like surface similarity in that it is used to describe perceptually evident properties of objects. Relational similarity refers to similarity in the relations between objects or propositions (Medin, et al., 1993). Gentner argues that relational similarity is central to analogy: Analogy is "a way of noticing relational commonalities independently of the objects to which those relations apply" (Gentner, 1988, p. 48).

The forced choice triad task method has been used to study novice-expert differences in similarity judgments in a variety of domains. Forced choice triad tasks require participants to make similarity judgments. Olson and Rueter (1987) suggest that indirect methods of knowledge elicitation such as similarity judgments are less vulnerable to reactivity effects than direct methods such as protocol analysis since they do not force the verbalization of non-verbal tasks. Ruiz-Primo and Shavelson (1996) and Medin (1989) support the use of indirect methods on the grounds that people cannot report their knowledge representations veridically. Indirect methods such as similarity judgments may also be less likely than methods such as interviews to induce researcher and social desirability biases, since the participant does not necessarily interact with the researcher while completing the task. Hardiman et al. (1989) suggest that similarity judgments are more discriminating than the sorting tasks used to investigate concept formation, since the former entails individual items that can be systematically modified whereas the latter requires participants to categorize a set of items under a single category name.

A Review of Reports on the Use of the Forced Choice Triad Task to Study Novice-Expert Differences

Hardiman et al. (1989), in a study that compared novice-expert differences in judging the similarity between classical mechanics problems using a forced choice triad task, found that both the novice and expert cohorts were influenced by the combination of surface and deep structures in a given problem. Each triad task trial consisted of three written descriptions of mechanics problems similar to those found in introductory textbooks. No diagrams were provided. The trials were designed based on four different similarity conditions: a referent item could match the target item on the basis of surface features only (S), deep features only (D), both surface and deep features (SD) and neither surface nor deep features (N). Therefore, there were four comparison types possible: S-D, S-SD, N-D and N-SD. It was hypothesized that the novice and experts would show different patterns of problem categorization: If experts base categorization exclusively on deep features and novices on surface features, experts would be expected to consistently select the item containing a deep comparison feature and novices would be expected to consistently choose the item containing a surface comparison feature. There were 32 items in total: two “model” problems were developed to reflect principles in each of energy, momentum, angular momentum and Newton’s Law of Kinematics. The model problems were presented four times, each with one of the comparison type pairings. Thus, the investigators systematically varied the task items to test the study hypothesis. The novice cohort was classified as undergraduate students (n=45) who had achieved a grade of at least B in an introductory course in classical mechanics; the expert cohort were eight PhD physicists and two graduate students nearing completion of PhD studies in physics. The authors offer no justification for the sample size, nor for the criteria by which they defined the novice and expert cohorts. The responses were scored and subjected to a three-factor analysis of variance with level of expertise, model problem type, and comparison type pairings as factors. Although the analysis showed that experts chose the item containing a deep feature significantly more often than did novices, the type of comparison pairing was found to influence the responses of both groups. The novice cohort did not match exclusively on surface features and when there was no surface item; i.e., the N-D items, they matched based on the deep feature to a much greater extent than predicted. In instances in which both referent problems had surface features that were similar to the target problem, experts were less likely to match based on deep similarities than in other conditions. The authors ascribed these findings to a posited ”similarity threshold”: Subjects are more likely to match based on an initial impression of surface similarity but if this option is unavailable, they are forced to consider the basis of their comparison more carefully and will be more likely to match based on deep features. These findings corroborate Tversky's (1977) contention that the objects in a stimulus set influence similarity judgments. The investigators found a correlation of r=0.3 between the novice cohort's performance on the forced choice triad task score and their performance on a separate test of solving mechanics problems. However, the authors note that a limitation of this experiment was that the basis on which similarity judgments were made were inferred rather than explicitly elicited. A second experiment was conducted in which novices were asked to compare a series of two item problems, indicate whether they would be solved similarly, indicate the reasons for their decisions and to complete a separate, problem solving test. This experiment corroborated the finding of a correlation between problem-solving ability and the tendency to perceive domain principles rather than surface features in problem-solving.

Rabinowitz and Hogan (2002) used a forced choice triad task to study novice-expert differences in categorizing statistics problems. These investigators also systematically varied the triad items in order to test their hypothesis by developing three different comparison types. In the first comparison type, one referent item was similar to the target item on the basis of only surface features such as characters and identity of variables, while the other referent item was similar only on the basis of structural features such as the nature of the variables, e.g., continuous versus nominal. In the second comparison type, one referent item had no similarities to the target item while the second referent item had only structural similarities to it. One referent item in the third comparison type had only surface similarities to the target item; the second referent item had no similarities to it. The participants were n=101 graduate students. The investigators divided the participants into three levels of experience based on the number of statistics courses they had taken. The novice group included those who had taken one or no statistics course, the intermediate group included those who had taken two or three courses, and the advanced group consisted of those who had taken at least four courses. They conducted a two-factor analysis of variance with experience level and comparison type as factors. A significant correlation of r=0.39 was found between level of experience and matching on the basis of structural features for comparison types one and two; no such significant correlation was found for comparison type three, supporting Hardiman et al's contention that novice and experts alike will match based on an initial impression of similarity surface.

Johnson and Mervis (1997) used the forced choice triad task method in their study of hypothesized novice-expert differences in basic and subordinate categorization of songbirds. Each trial consisted of silhouettes of three songbirds; two of the same species and one of a different genus chosen on the basis of its high similarity to the same-species birds in order to make the task of differentiating between them more difficult. In all cases, the birds were shown against a white background facing to the right. Advanced (n=12) and intermediate (n=20) expert cohorts were identified on the basis of peer ratings of their knowledge. The novice group was undergraduate psychology students (n=28) who reported low knowledge about songbirds. The authors offer no justification for their sample size. Although they indicate that peer ratings were made on a seven-point scale, they do not indicate what rating differentiated the advanced from the intermediate expert cohort. Another criticism that might be made of this study is that participants were primed by a practice forced choice triad task, in which they were not allowed to proceed to the next trial unless they matched on the basis of a taxonomic feature. Responses were scored as “taxonomic” if the same-species birds were matched and “non-taxonomic” if the distractor bird was matched. The responses were subjected to a one-way analysis of variance, with knowledge level as the between-subjects factor. No significant difference was found in scores between the advanced and intermediate expert groups; however, novices scored lower than both expert cohorts on the number of their taxonomic responses. The investigators also recorded and analyzed response times; they found that the novice cohort’s response times were significantly faster than those of the two expert cohorts. They ascribed this difference to a greater degree of deliberation on the part of the expert cohort.

Rabinowitz and Hogan (2009) investigated novice-expert differences in teachers’ problem representations using a forced choice triad task method. They developed 18 trials consisting of descriptions of typical classroom problems; the referent items that contained surface features provided information describing context or setting, such as classroom subject or grade, while the items that contained deep features reflected pedagogical theory in the areas of assessment, instruction, and equity. Scenarios were randomly assigned to a trial as either a referent or target item. These investigators systematically varied the items by designing the scenarios to reflect one of three problem types to determine whether the combination of features influenced the problem representation. In the first problem type, one referent item had a deep similarity to the target item while the second referent item had a surface similarity to it. In the second problem type, one referent item had a surface similarity to the target item while the other referent item had neither surface nor deep similarities to it. The third problem type consisted of one referent item with a deep similarity to the target item and one referent item that had no similarity to the target item. Experienced teachers reviewed the scenarios and modifications were made based on their review. The novice cohort consisted of n=98 undergraduate students from a college teacher preparation program. The expert cohort consisted of n=20 National Board for Professional Teaching Standards certified teachers. The validity of this certification was supported by study findings on a comparison between certified and non-certified teachers on 13 attributes derived from the literature on teaching expertise. The responses on the forced choice triad task were scored on the basis of problem type: For problems in which the referent item having a deep similarity to the target was chosen, a score of 1 was given. For problems in which the referent item having a surface similarity was chosen, a score of 1 was given if the other referent item had no similarity to the target. The scores were analyzed using a two-factor analysis of variance with comparison type as the within subjects factor and practice level as the between-subjects factor. The analysis showed significant main effects of practice level and problem type but no interaction effect.

Devantier et al. (2008) used a forced choice triad task to evaluate novice-intermediate-expert differences in patient diagnosis and medical management. Subjects were presented with a series of ten trials, half of which related to management and half to diagnosis. The diagnostic cases related to diabetes, while the management cases were reported to be more general in nature. While the authors did not explicitly indicate the comparison types used in their trials, it is assumed from the example they provided that, for each trial, one referent item was designed to be similar to the target item based on surface similarities and the other on the basis of deep similarities. Practice level of participants was categorized on the basis of academic training and clinical specialization. The novice group consisted of n=15 second and third year medical students who responded to a recruitment email; the intermediate cohort consisted of n=8 residents in the first or second postgraduate year who were recruited by the authors; and the expert cohort was n=13 endocrinologists who were recruited from across Canada via email by one of the authors. The clinical specialties of the medical residents were not reported. As in the other studies, the authors did not report the rationale for their sample size. Unlike some of the other studies reviewed, these investigators asked participants to provide a rationale for their similarity judgments so that scoring was not solely dependent on their inferences. The investigators scored the responses as reflecting a match based on surface or deep features; the responses were also scored by two independent raters who were blinded to practice level. The scores were recorded as a proportion of deep to total responses, and subjected to one-way analysis of variance with practice level as the between-subjects factor. Overall, the expert cohort had higher scores (indicating matching on a deeper level of similarity) than the novice cohort; there was no significant difference in scores between the intermediate and expert cohorts. Analyses were conducted of performance on the management and diagnosis trials separately. For the management cases, the expert and intermediate cohorts were similar in matching based on deep features, while the novice group were found to match predominantly on the basis of surface similarities. For the diagnostic cases, experts again matched primarily on the basis of deep similarities, while the intermediate and novice groups were more likely to match based on surface features. The authors suggest that more experience with management than with diagnosis at the intermediate level might account for these findings.

Many of the therapeutic interventions used to treat cardiovascular disorders entail manipulation of the determinants of cardiac output[3]: preload[4], afterload[5], and contractility[6]. Cardiovascular specialty nursing practice, therefore, requires a sound understanding of hemodynamic[7] principles. In the present study, it was hypothesized that nurses nearer the expert end of the novice-expert continuum would make similarity judgments based on deep, structural relational similarities, such as those based on hemodynamic principles, to a greater degree than would those nearer the novice end of

[3] The volume of blood ejected by each ventricle per minute; cardiac output equals heart rate multiplied by stroke volume (the volume of blood ejected by the ventricle in systole).
[4] The degree of myocardial fibre stretch at end-diastole; according to Frank-Starling's Law of the Heart, stroke volume (the volume of blood ejected from the ventricles during systole) increases as preload increases, up to a physiological limit. Preload cannot be measured directly; in the clinical setting it is inferred from atrial end-diastolic pressure measurements.
[5] Resistance to ventricular ejection in systole; the chief determinant of left ventricular afterload is the degree of systemic arteriolar constriction. The higher afterload is, the more pressure the left ventricle must generate to eject stroke volume and the greater the myocardial oxygen consumption (MVO_2). In the setting of myocardial infarction, the primary therapeutic goal is to minimize the imbalance between myocardial oxygen supply and demand.
[6] The force of myocardial contraction during systole; as contractility increases, stroke volume increases.
[7] Literally, the movement of blood through the closed circuit of the cardiovascular system.

the continuum. A forced choice triad task was developed and tested to investigate this hypothesis. Each trial in the task consisted of three items: a “target” item and two “referent” items. For each trial in the task, participants were asked to select the referent item that they believed was more similar to the target item and to indicate the reason for their choice. Two comparison types of items were developed. In the surface-deep (SD) comparison type of trial, one referent item was designed to have surface or superficial characteristics, such as patient age, sex, or diagnosis, that were similar to the target item, while the other referent item was designed to be similar to the target item on the basis of deep, structural or relational features reflecting the principles of hemodynamics. In the surface-deep/surface (SDS) comparison type, one referent item was designed to be similar to the target item on the basis of both surface and deep features while the other referent item was designed to be similar to the target based on surface features only.

Instrument Reliability and Validity

As in other methods used to study novice-expert differences, key issues in the use of the forced triad task method are the criteria by which cohorts are defined and the validity of the instrument in discriminating differences between them. The problem of classifying practice level might be avoided by studying individuals at increments along the novice-expert continuum; thus, a novice/less experienced cohort can be identified on the basis of lacking the practice longevity necessary for expertise while an experienced/competent cohort can be defined by criteria such as advanced educational attainment, practice longevity and clinical specialization. The latter cohort could reasonably be expected to perform more like experts than the novice/less experienced cohort on a forced choice triad task designed to test this study's hypothesis.

The reliability of an instrument refers to the consistency of repeated measurements on it (Carmines & Zellers, 1979). The retest method of evaluating reliability entails administering the test to the same people after a period of time has elapsed (Carmines & Zellers, 1979). Carmines and Zellers (1979) assert that the retest method is subject to limitations, including the misinterpretation of change in the construct being measured as measurement instability, reactivity, and the overestimation of an instrument's reliability due to the influence of memory. An alternative-form method of reliability assessment entails administering two tests measuring the same construct to the same individuals (Carmines & Zellers, 1979). The main limitation of this method is the difficulty of designing two tests that are sufficiently parallel (Carmines & Zellers, 1979). The internal consistency method of assessing reliability entails evaluating the extent to which a test's items measure the same construct (Carmines & Zellers, 1979). The reliability coefficient, Cronbach's alpha, is a measure of internal consistency and ranges from 0.0 to 1.0; a value of alpha of 0.70 to 0.90 reflects a good to very good level of internal consistency. The point-biserial (PBS) correlation provides an indication of item discrimination or a measure of association between an individual's performance on each item on a test and on the remaining items on the test (Norman & Streiner, 2008). The value of the point-biserial correlation is from -1.0 to +1.0, with higher values reflecting a higher degree of consistency and discrimination A point-biserial correlation of > 0.3

reflects a good or very good degree of correlation between an item and overall test difficulty (University of Texas at Austin, n.d.).

Criterion validity refers to the extent to which scores on an instrument correlate with scores on a criterion variable that measures the same construct, or a “gold standard” (Carmines & Zellers, 1979; Roach, 2006). Concurrent validity is evaluated when scores or measurements on the instrument under consideration and those on the criterion are obtained at the same time and compared (Carmines & Zellers, 1979). Predictive validity is evaluated when scores on the instrument are obtained and compared to scores subsequently obtained on the criterion variable (Carmines & Zellers, 1979). Carmines and Zellers (1979) point out that for many of the phenomena of interest in the social sciences, no criterion variable exists. Since there is no known “gold standard” for nurses’ similarity judgments, the criterion validity of participants’ scores on the instrument used in this study cannot be evaluated.

Cronbach and Meehl (1955) state:

> Construct validity is ordinarily studied when the tester has no definite criterion measure of the quality with which he is concerned, and must use indirect measures. Here the trait or quality underlying the test is of central importance, rather than either the test behavior or the scores on the criteria (p.282).

Evidence of construct validity is accrued when an empirical measure of a construct is consistent with a theoretically-derived prediction (Carmines & Zellers, 1979). Carmines and Zellers (1979) suggest three possible reasons that observed data may not support a hypothesized prediction: 1) the instrument does not measure the construct of interest; 2) the theory used to generate the hypothesis or prediction is incorrect; or 3) the method (including the statistical analysis) used to test the prediction or hypothesis is inappropriate.

Lynn (1986) asserts that validating the content of an instrument entails a two-stage process: a developmental stage and a judgment-quantification stage. In the developmental stage, the researcher seeks to identify the full content domain that the instrument is intended to measure, and then attempts to generate and refine instrument items that reflect this content (Lynn, 1986). The judgment-quantification stage requires the adjudication by content matter experts as to how well instrument items represent specific areas of content domain and of the comprehensiveness of the entire instrument in representing the domain (Lynn, 1986).

Summary

The investigation of novice-expert differences in memory and cognition within the domain of specialty cardiovascular nursing could provide evidence to inform pedagogical strategies. Novice-expert differences in cognition and problem-solving have been attributed to differences in the organization of knowledge in memory. Retrieval of knowledge from memory entails a judgment of the similarity between a current situation or problem state and one encoded in memory. Given the prominence of the construct of similarity in cognition and memory, a similarity judgment task, such as a forced choice triad task, seems particularity appropriate for the investigation of novice-expert differences. This method avoids some of the potential sources of bias observed in other methods. Furthermore, the comparison types within items comprising the task can be varied systematically to enable a secondary analysis of methodological artefact.

The study of hypothesized differences at increments along the novice-expert continuum avoids some of the difficulties associated with classification of practice level. This study used a forced choice triad task method to investigate hypothesized differences between cardiovascular nurses' perception of domain-specific clinical scenarios at two increments on the novice-expert continuum.

Chapter Two

Specialized knowledge and skills underlie the safe nursing care of patients with cardiovascular disease (Canadian Council of Cardiovascular Nurses, 2008). Specialty nursing education, including cardiovascular nursing, has little empirical foundation, in part, because of a paucity of research. The study of differences between novice and experts in a domain may yield findings that could inform pedagogical strategies for specialty nursing education. In a variety of domains, novice-expert differences in perception have been shown, with novices perceiving primarily the perceptually obvious or surface features of scenarios and experts perceiving predominantly the deep features of scenarios that reflect the domain’s principles. The forced choice triad task method is useful for testing the hypothesis of the existence of these differences since it entails a judgment of the similarity between items designed to demonstrate deep and surface features relevant to the domain of interest.

In this chapter, the research purpose, questions and hypothesis are identified and the phases of the development and testing of the forced choice triad task items used in the study are described.

Study Purpose, Research Question and Hypothesis

It has been posited that there are differences across the novice-expert continuum in the perception of domain-specific features, with those nearer the novice end of the continuum perceiving primarily superficial, surface, or perceptually obvious features and those nearer the expert end of the continuum perceiving to a greater extent the deep, structural principles underlying the domain (Chi et al., 1981; Devantier et al. , 2008; Hmelo-Silver et al., 2007; Hmelo-Silver & Pfeffer, 2004; Rabinowitz & Hogan, 2002). Many of the interventions used to treat cardiovascular disorders entail manipulating hemodynamics. Cardiovascular nurses are responsible for administering or monitoring many of these interventions. Therefore, cardiovascular nurses should understand the principles of hemodynamics. If, as hypothesized, a more experienced cohort of nurses perceives the deep principles underlying a domain-relevant scenario to a greater degree than a less experienced cohort, the more experienced cohort should make similarity judgments on the basis of deep domain features, such as the principles of hemodynamics, to a greater degree than the less experienced cohort.

Purpose of the Study

The purpose of this study was to investigate hypothesized differences in the similarity judgments of nurses at two points in the novice-expert continuum of cardiovascular nursing practice.

Primary Question:

Are there differences in the nature of the similarity judgments between novice/less experienced and experienced competent nurses?

Null Hypothesis: Experienced/competent nurses will not make a higher proportion of similarity judgments based on the principles of hemodynamics in a forced choice triad task than novice/less experienced nurses.

Alternative Hypothesis: Experienced/competent nurses will make a higher proportion of similarity judgments based on the principles of hemodynamics in a force choice triad task than novice/less experienced nurses.

Methods

Sample Criteria

Because practice longevity is a necessary but insufficient condition for the development of expertise (Bobay, 2004), and because practice level--particularly at the competent level and beyond--is difficult to designate, the sample for this study was drawn from among novice/less experienced and experienced competent practitioners. The novice/less experienced cohort was defined as registered nurses who have worked less than three years in dedicated cardiac units. The experienced/competent cohort was identified on the basis of practice longevity, advanced academic qualifications, and clinical specialization: Nurse Practitioners with at least ten years of experience in cardiac nursing were defined as experienced/competent practitioners. In Ontario, Nurse Practitioners are prepared at the master`s degree level (Nurse Practitioners` Association of Ontario, n.d.), are required to complete 700 hours of clinical training in their program, and must pass an Objective Structured Clinical Examination (OSCE) at two points in their training (University of Toronto, n.d.). The criteria defining the two cohorts are summarized in Table 1 below.

Table 1 *Criteria Used to Define Practice Cohorts*

Cohort	Designation	Highest Level of Education	Practice Longevity in Cardiovascular Nursing[8]	Employment
Novice/less experienced	Registered Nurse	BScN or Diploma	Less than three years	Dedicated adult cardiac inpatient units: cardiology, cardiac surgery, coronary care, cardiovascular intensive care
Experienced/competent	Nurse practitioner	MScN	At least ten years	Dedicated adult cardiac inpatient units: cardiology, cardiac surgery, coronary care, cardiovascular intensive care

Sample Size Calculation

Lenthe (2001) cautions against the use of "canned" effect sizes in the calculation of sample size. He argues that calculating sample size by using a small, medium or large effect size at a given power disregards study elements that influence effect size (Lenthe,

[8] For the competent/experienced cohort, practice longevity includes all experience in cardiovascular nursing, including experience as a registered nurse before becoming a nurse practitioner.

2001). The t-test will be used to compare the means scores of the two cohorts, since the independent or predictor variable in this study is a nominal variable (practice level) and the dependent variable is a continuous variable (score on the task). The required sample size for this study was calculated based on the findings of a similar study by Devantier et al. (2008). Using a forced choice triad task, these authors investigated novice-expert differences in physicians’ perceptions of patient care scenarios: On average, the expert cohort (n=11) based their similarity judgments on deep features 50% of the time whereas the novice cohort (n=15) did so 30% of the time (Devantier, et al., 2008). Therefore, a mean difference of d=0.2 in the proportion of similarity judgments based on deep features was found between the expert and novice cohorts. The largest standard error of the mean was found within the expert cohort (SEM=0.07); assuming that a similar difference would be found between cohorts in the present study, this value was used to calculate the sample size (Browner, Newman, & Hulley, 2007 pp.67-68).

$$\because \text{SEM} = \frac{\text{standard deviation (s)}}{\sqrt{\text{n}}}$$
$$\therefore 0.07 = \frac{s}{\sqrt{11}}$$
$$\therefore 0.07 = \frac{s}{3.32}$$
$$\therefore \text{s} = 0.07(3.32)$$
$$= 0.23$$

$$\text{Sample size} \quad = \frac{16(\text{Standardized effect size})^2}{\text{d}^2}$$
$$= \frac{16(0.23)^2}{0.2^2}$$
$$= \frac{16(0.053)}{0.04}$$
$$= \frac{0.85}{0.04} = 21.25$$

Figure 1 Sample Size Calculation

Therefore, a sample size of n=21 in each cohort was required for a total sample size of n=42.

Recruitment

A snowball sampling technique was initially used to recruit participants for the study. Snowball sampling is a type of convenience sampling, in which early participants are asked to refer others who meet eligibility criteria (Polit, Beck, & Hungler, 2001). Snowball sampling is used to recruit participants who belong to groups that are difficult to recruit by other means because they have very specific traits or because they belong to "hidden" or marginalized groups (van Meter, 1990). The eligibility criteria for this study were quite specific as to practice longevity and area of practice. The College of Nurses of Ontario (CNO) maintains a mailing list of registrants who have consented to be contacted for research purposes; researchers select the demographic and practice characteristics of interest when requesting a mailing list. However, this information was not sufficiently specific for the purposes of this study, since statistics collected are not classified by practice specialty. A very large mailing using the CNO's mailing list requiring nurses to self-report their eligibility was thought to be economically infeasible. Therefore, a snowball sampling technique was initially used, whereby early sample members known in a professional capacity by the writer and by supervisory committee members to meet eligibility requirements were asked to identify other eligible potential participants. When it appeared that the snowball sampling technique would not generate a sufficiently large sample, application for ethics amendment approval was sought and obtained from the

McMaster Student Research Ethics Committee to put up posters at all sites of Hamilton Health Sciences and to meet with groups of nurses for recruitment purposes.

A modified Dillman technique was used to distribute the study instrument. This method, originally designed for surveys and also known as the tailored design method (TDM), has been demonstrated to yield superior response rates in comparison with other mailing protocols (Dillman, 2007). In the Dillman method, four contacts are made via first-class mail, with an additional contact via telephone or certified mail; stamped return envelopes are included with mailouts of personalized correspondence and study instruments, and financial incentives are provided (Dillman, 2007). Due to time and economic constraints, this protocol was modified for this study. Initial approach for some potential participants was via email (Appendix A); those who agreed to participate were mailed via post a personalized introductory letter (Appendix B), a consent and information sheet (Appendix C), a demographic questionnaire (Appendix D) , the study instrument (Appendix E), a response sheet (Appendix F) and a stamped, return envelope. If, after two weeks, the completed demographic questionnaire and study instrument were not returned, a reminder letter with the study materials was sent. Upon receipt of the completed study tools, participants were mailed a thank you letter (Appendix G), a Tim Horton’s gift card in the amount of $10, and a copy of their signed consent form for their own records.

Ethical Issues

Participants may have been concerned that they lacked sufficient knowledge to perform the task, and as a result, they may have been fearful of embarrassment. They may also have been concerned that perceived poor performance on the task may have negative implications for their licensure or employment status. To mitigate these concerns, it was explicitly stated in the study consent/information sheet that 1) there are no correct or incorrect responses to the series of tasks; 2) responses are confidential and results will be reported as aggregate data; and 3) participation in the study would in no way affect licensure or employment status. Each participant was assigned a code number that was used to anonymize him or her and to assist in cross-reference mailing materials. Electronic data files were password protected.

Approval to conduct the study was obtained from the McMaster University Student Research Committee in July, 2009; this approval allowed recruitment of participants from all sites of Hamilton Health Sciences Centre. The investigator enquired as to whether institutional research ethics board approval was required to recruit participants from organizations in which eligible participants might be employed: Two community hospitals and one academic health science centre did not require such approval. Research ethics board approval was required and obtained to recruit staff members at a second academic health science centre. A third academic health sciences centre required institutional research ethics board approval to recruit its staff members;

however, since the adjudication process for this approval was stated to require four to six months, the application was not made.

Instrument Development

Phase 1.

Development of the prototype instrument.

The first phase entailed the development of an initial series of forced choice triad trial items to test the study hypothesis. Each trial consisted of three brief items: a “target” item labelled “A” and two “referent” items, labelled “B” and “C”. In each triad, one referent item was designed to have superficial characteristics, such as patient age and sex or diagnosis, that were similar to the target item, while the other referent item was designed to be similar to the target item on the basis of deep features. The deep similarities between referent and target items represent three categories that reflect principles of hemodynamics: 1) manipulation of the determinants of cardiac output (preload, contractility, afterload and heart rate) to achieve a therapeutic goal; 2) the operation of cardiovascular compensatory mechanisms to maintain homeostasis (such as reflex increase in heart rate to compensate for a decline in stroke volume); and 3) the abnormality in the determinants of cardiac output that underlies a particular pathophysiology (such as excessively low afterload as the basic pathophysiologic feature of certain shock states). The basis on which each of the original ten forced choice triad trials was developed in the original instrument is shown in Table 2.

Table 2 *The Basis of Surface and Deep Similarities of Referent Items to Target Item for Each Forced Choice Triad Task in Original Instrument*

Task	Referent Item B Surface Similarities to Target	Referent B Deep Similarities to Target	Basis of Deep Similarity	Referent C Surface Similarities to Target	Referent B Deep Similarities to Target	Basis of Deep Similarity
1	Age, sex, diagnosis, mention of drug therapy			Mention of drug therapy	Drug therapy to reduce preload	Therapeutic goal
2	Hypotension	Excessively low afterload	Basic physiologic problem	Age. Hypotension Cardiac problem		
3	Abnormal vital signs Age	Operation of compensatory mechanisms: tachycardia and increased diastolic BP	Operation of compensatory mechanisms	Age Cardiac problem Tachycardia		
4	Arrhythmia	Manipulation of heart rhythm/rate to improve cardiac output	Therapeutic goal	Age, sex, cardiac problem		

Table 2 (Continued)

Task	Referent Item B Surface Similarities to Target	Referent B Deep Similarities to Target	Basis of Deep Similarity	Referent C Surface Similarities to Target	Referent C Deep Similarities to Target	Basis of Deep Similarity
5	Coldness, pallor of extremities			Coldness, pallor	Selective vasocon-striction to maintain vital organ perfusion	Operation of compensa-tory mechanisms
6	Cardiac problem; mention of drug therapy	Manipulation of preload to improve cardiac output	Therapeutic goal	Diagnosis; mention of drug therapy		
7		High afterload	Basic physiologic problem	Age; valvular problem		
8	Age; valvular surgery				Manipulation of heart rate or rhythm to improve cardiac output	Therapeutic goal
9		Increased resistance to diastolic filling	Basic physiologic problem	Age; sex; cardiac surgery		
10	Shock state				Excessively low afterload	Basic physiologic problem

In the example below, referent item B was designed to be similar to target item A on the basis of the superficial or surface features of patient age and the mention of drug therapy. Referent item C was designed to be similar to target item A on the basis of both superficial features (mention of drug therapy) and deep features (pharmacologic reduction of preload as the therapeutic goal of treatment).

MK is a 67-year-old man with congestive heart failure. He receives Lasix, with the dosage based on his daily weight.

A

EM is a 69-year-old man who developed heart failure after a large anterior wall MI. He receives ramapril 7.5 mg po once daily.

B

JH is a 55-year-old woman with stable angina. She receives nitroglycerin 0.3 mg S/L prn for chest pain.

C

Figure 2 Forced Choice Trial One in Original Instrument

In the trial below, referent item B was designed to be similar to target item A on the basis of a superficial feature (hypotension) and a deep feature (excessively low afterload as the basic pathophysiologic problem), while referent item C was designed to be similar to target item A based on surface features of the patient's age, the presence of hypotension and of a cardiac problem.

KF is a 33-year old woman who developed septic shock after mitral valve replacement surgery. Her blood pressure is 80/46 mm Hg.
A

JL is a 60-year-old man with a severe peanut allergy. He unknowingly ate food containing peanut oil in a restaurant. His blood pressure is 77/40 mm Hg.
B

OP is a 40-year-old man who developed congestive heart failure after a large anterior wall MI. He received Lasix 40 mg IV and diuresed 2 litres in six hours. His BP is 86/46 mm Hg.
C

Figure 3 Forced Choice Trial Two in Original Instrument

For each trial, participants were asked to select the referent item (B or C) that they believed was more similar to the target item (A), and to indicate the reason for their choice. Participants were originally asked to judge similarity based on only one feature to preclude the possibility of choosing referent items on the basis of both surface and deep features for the same trial. Since this instruction was ignored by some participants, it was decided to score responses according to the "highest" level of response; i.e., if the respondent matched items based on both superficial and deep features, the response was scored as a deep match, since the respondent demonstrated the ability to match based on deep featural similarity for that trial. Additionally, if participants justified their matching with incorrect information, the response was scored as a surface match.

The prototype instrument consisting of ten trials was tested with n=2 nurses; one in the novice/less experienced and one in the experienced/ competent cohort. Each trial was scored by the primary researcher as to whether similarity judgments were made based on superficial or deep similarities between the chosen referent item and the target item. Phase 1 testing also entailed an assessment of clarity of instructions, clarity of case descriptions, and length of time required for completion (Appendix H). Both participants in Phase 1 of testing indicated that the instructions for completing the task were clear. One participant indicated that the fifth and sixth trials should be revised but did not indicate the reason for this. The first participant indicated that the instrument took 22 minutes to complete, while the second participant indicated that it took 10 minutes to complete. On the basis of participant matches, some forced triad task items did not seem to be sufficiently clear. Some items contained irrelevant information, such as the time of day, while some items contained more information and more detailed descriptions than others. Therefore, existing items were revised and six additional items were developed.

Phase 2.

Evaluation of instrument clarity, representativeness and comprehensiveness.

A panel of four content experts was recruited to evaluate the revised study instrument. Three of the content experts were prepared at the master's degree level in nursing; one was doctorally prepared. Three of the content experts were clinical nurse specialists in a variety of cardiac subspecialties: one in heart failure/ventricular assist devices/transplant, one in interventional cardiology and arrhythmia management, and one in cardiology and interventional cardiology. The fourth content expert was a clinical educator in cardiology and cardiac surgery.

Sixteen triad tasks (Appendix I) were evaluated by the context experts for clarity, comprehensiveness and representativeness. For each of the target and referent items within a trial, similar information was provided and parallel language was used. Content experts independently rated on a four-point scale each referent item as to how well it reflected the stated basis of the surface or deep match (Appendix J). Content experts were provided with a chart showing the basis for deep and surface matches on which the items were designed (Table 3).

Table 3 *The Basis of Surface and Deep Similarities of Referent Items to Target Item for Each Forced Choice Triad Task in Revised Instrument*

Task	Referent B Surface Similarities to Target	Referent B Deep Similarities to Target	Basis of Deep Similarity	Referent C Surface Similarities to Target	Referent C Deep Similarities to Target	Basis of Deep Similarity
1	Age; sex; diagnosis; mention of drug therapy			Mention of drug therapy	Drug therapy to reduce preload	Therapeutic goal
2	Hypotension; age; shock state	Excessively low afterload	Basis patho-physiology	Hypotension; cardiac problem; shock state		
3	Abnormal vital signs; Age	Tachycardia; increased diastolic BP	Operation of compensatory mechanisms	Age; cardiac problem; tachycardia		
4	Arrhythmia	Manipulation of heart rhythm/rate to improve cardiac output	Therapeutic goal	Age; sex; cardiac problem		
5	Coldness, pallor of extremities; age			Coldness, pallor of extremities	Selective vaso-constriction to maintain vital organ perfusion	Operation of compensatory mechanisms
6	Hypertension				High RV afterload	Basic patho-physiology

Table 3 (Continued)

Task	Referent B Surface Similarities to Target	Referent B Deep Similarities to Target	Basis of Deep Similarity	Referent C Surface Similarities to Target	Referent C Deep Similarities to Target	Basis of Deep Similarity
7	Valve surgery				Control of rhythm to maintain BP	Therapeutic goal
8	Age; valvular surgery				Manipulation of heart rate or rhythm to improve cardiac output	Therapeutic goal
9		Increased resistance to diastolic filling	Basic patho-physiology	Age; sex; cardiac surgery		
10	Shock state			Shock state	Excessively low afterload	Basic patho-physiology
11	Pulmonary problem	High RV afterload	Basic patho-physiology	Valvular stenosis		
12	Drug therapy			Drug therapy	Therapy to reduce afterload	Therapeutic goal
13	Age; valvular surgery				Volume expansion to maintain CO	Therapeutic goal

Table 3 (Continued)

Task	Referent B Surface Similarities to Target	Referent B Deep Similarities to Target	Basis of Deep Similarity	Referent C Surface Similarities to Target	Referent C Deep Similarities to Target	Basis of Deep Similarity
14	Valvular stenosis				Conditional that may cause right-sided failure	Basic patho-physiology
15		Heart rate excessively low	Basic patho-physiology	Pacemaker; sensing problems		
16	Shock states				Conditions resulting in high CO	Basic patho-physiology

Items that were scored as not reflecting or reflecting poorly the stated basis for matching were discarded; that is, those items that were scored as 1 or 2 on the scale by at least one content expert were discarded (Table 4).

Table 4 *Content Experts' Ratings of Stated Basis of Similarities of Referent Items to Target Items for Each Forced Choice Triad Task*

Item	Ratings 1	2	3	4	Action
1. B	0	0	0	4	Retain
1. C	0	0	0	4	Retain
2. B	0	0	0	4	Retain
2. C	0	0	0	4	Retain
3. B	0	1	2	1	Discard
3. C.	0	1	0	3	Discard
4. B	0	0	3	1	Revise
4. C	0	0	1	3	Retain
5. B.	0	0	0	3	Retain
5. C	0	0	1	2	Retain
6. B	0	1	1	2	Discard
6. C	0	1	1	2	Discard
7.B	0	0	1	3	Retain
7. C	0	0	1	3	Retain
8.B	0	0	0	3	Retain
8. C	0	0	0	3	Retain
9. B	0	1	0	3	Discard
9. C.	0	0	1	3	Retain
10. B.	0	0	0	4	Retain
10. C	0	1	0	3	Discard
11. B	0	0	0	4	Retain
11. C	0	0	0	4	Retain
12. B	0	0	1	3	Retain
12. C	0	0	0	4	Retain
13. B	0	0	0	4	Retain
13. C	0	0	0	4	Retain
14. B	0	0	0	4	Retain
14. C	0	1	1	2	Discard
15. B	0	0	1	3	Retain
15. C	0	0	0	4	Retain
16. B	0	0	0	4	Retain
16. C	0	0	0	4	Retain

Six trials (trials 3, 5, 6, 9, 10, 14) were discarded on the basis of this scoring. Content experts were asked to identify concepts important to nurses' understanding of hemodynamics that were not represented in the tasks; no additional concepts were identified.

Phase 3.

Testing of revised instrument.

The revised instrument was tested with n=4 registered nurses; n=2 in each practice level cohort. Phase 3 testing entailed an assessment of clarity of instructions, clarity of case descriptions, and length of time required for completion. Each trial was scored by the primary researcher as to whether similarity judgments were made based on a superficial or a deep feature. Scoring was also done by a second, independent rater who was blinded to the practice level of participants. A score of 1 was assigned for a match based on a deep similarity between referent and target items, and a score of 0 was assigned for matching based on surface features. Disagreements about scoring were resolved through discussion between raters. A point-biserial correlation analysis was conducted, in which participants' performance on each trial was correlated with their performance on the other trials in the instrument. This analysis provides a measure of discrimination and, therefore, an indication of internal consistency. The point-biserial calculation was done in PASW 17.0 and results are shown in Table 5. Cronbach's alpha, an indicator of internal consistency, was also calculated using PASW 17.0**.**

Table 5 *Point-Biserial Correlation for Items in Revised Study Instrument.*

Item	1	2	3	4	5	6	7	8	9	10
Point-biserial correlation	0.29	0.98	0.33	0.57	0.33	0.29	0.29	0.50	0.98	0.50

Cronbach's alpha was calculated at α= 0.80. An alpha level of 0.70 to 0.90 and a point-biserial correlation of greater than 0.3 reflect an acceptable level of internal consistency. No trials were discarded on the basis of the calculated values of the point-biserial correlation or Cronbach's alpha.

Phase 4.

Results.

The final instrument distributed to study participants consisted of 10 trials. Because the study instrument was not revised on the basis of the phase 3 testing, the scores from the phase 3 testing were included in the final analysis for a total of n=17. Demographic characteristics of the study participants are summarized in Table 6.

Table 6 *Demographic Characteristics of Study Participants*

		Practice Cohort	
		Novice/Less Experienced	Competent/Experienced
Unit			
	Cardiology	4	3
	Cardiac Surgery	3	6
	Cardiovascular Intensive Care Unit	1	0
	Coronary Care Unit	0	0
Education: Highest Attained	Diploma	1	0
	BScN	7	0
	MScN	0	9
	PhD	0	0

Scoring of responses was done by the primary researcher and by an independent rater who was blinded to practice level. Participants occasionally identified similarities between target and referent items that were not anticipated in the design of the items. Furthermore, participants sometimes matched target and referent items on ostensible deep similarities; however, their stated rationales for matching suggested a surface response. Disagreements between raters were resolved through discussion via email. There were a total of 19 disagreements between raters over scores on the ten trials for each of the 17 completed study instruments; this represents an 11% disagreement rate. Scores did not change appreciably when those trials on which raters did not initially agree were omitted from the analysis [novice-less/experienced M= 0.28 (0.36); competent/experienced M= 0.40 (0.24)].

Raters disagreed most frequently on scores of trials three and four (four disagreements on scores for trial three and six disagreements on scores for trial four). For example, for participant #5, the primary investigator scored the response on trial four as a surface response, while the independent rater scored it as a deep response. The primary investigator's rationale was "I called this a surface match because the participant has identified that these items are similar because they both involve valvular surgery; item A indicates that the bradycardia was post-cardiac surgery, so I don't think she gets any points for stating that." The independent rater's response was "I agree, they are both post-op. I was initially thinking that the nurse is aware of a potential complication of AVR, but could be quite content to follow your rationale and label the response as 'surface'". Raters disagreed on scores of trials three and four for participant #7. The primary investigator scored the response for trial three as surface and the independent rater scored it as a deep response. The primary investigator's rationale was "I called this a surface match because the participant based the match on the fact that both patients have arrhythmias". The independent rater replied, "I agree, the participant comments on arrhythmias but as least he/she commented about the possibility of being symptomatic with arrhythmia, so I gave 'deep'--evidence of 'thinking ahead'". In scoring trial four as a surface match for the same participant, the primary investigator commented "I called this a surface match because, again, the participant has only identified that both patients have an arrhythmia that is being treated. If she had gone a little farther and said that the goal of therapy is to optimize cardiac output by optimizing heart rate/rhythm, I might have called this a deep match." The independent rater concurred with this rationale. The proportion of deep similarity judgments was calculated for each individual. These data and the demographic data, stripped of identifying information, were entered into the PASW 17.0 statistical analysis software program. PASW was used to generate descriptive statistics: the number (n) of participants in each practice level, and the mean proportion of deep responses and standard deviation for the ten trials for each practice level cohort.

The independent t-test is used to test whether the mean of a continuous variable differs significantly between two groups (Browner et al., 2007). Assumptions underlying the use of the t-test are that the samples are drawn from a normally distributed population (Daniel, 2005) and have equal variances (Kinnear & Gray, 2009). The t-test is robust to violations of its statistical model, provided that the sample sizes are equal or nearly equal and are not too small. According to the Central Limit Theorem, a normal distribution of test scores can be assumed if the sample size is larger than 30 (Daniel, 2005). Although only 17 nurses participated in the study, the sample sizes were almost equal with n=9 in the competent/experienced cohort and n=8 in the novice/less experienced cohort. Furthermore, Levene`s test of homogeneity of variance was non-significant (p=0.17), indicating that an assumption of equal variance between test scores for the two cohorts could be made. No significant difference was found between the scores of the two study cohorts with respect to the number of deep matches made [t (0.95, 15) = - 0.89, p=0.19]. However, a difference in the hypothesized direction between scores was found (Table 7).

Table 7 *Percentage of Matches that Aligned with the Deep Comparator on the Forced Choice Triad Task for Novice/Less Experienced and Competent/Experienced Cohorts*

Practice level	M	SD	SEM
Novice/less experienced (n=8)	0.29	0.33	0.12
Competent/experienced (n=9)	0.41	0.24	0.08

These data were used to calculate a measure of effect size, d; the calculation uses the mean scores of two independent samples and their pooled standard deviations as an approximation of the population parameters, μ and σ, respectively (Figure 4).

$$s = \sqrt{\frac{SS_1 - SS_2}{df_1 - df_2}}$$
$$= \sqrt{\frac{(n_1 - 1) \times s_1^2 + (n_2 - 1) \times s_2^2}{n_1 + n_2 - 2}}$$
$$= \sqrt{\frac{[(8 - 1) \times 0.33^2] + [(9 - 1) \times 0.24^2]}{8 + 9 - 1}}$$
$$= \sqrt{\frac{[7 \times 0.11] + [8 \times 0.06]}{16}}$$
$$= \sqrt{\frac{0.77 + 0.48}{16}}$$
$$= \sqrt{\frac{1.25}{16}}$$
$$= \sqrt{0.08}$$
$$= 0.28$$

$$\because d = \frac{\mu_1 - \mu_2}{\sigma}$$
$$\therefore d = \frac{0.41 - 0.29}{0.28}$$
$$= \frac{0.12}{0.28}$$
$$= 0.43$$

Figure 4 Calculation of Effect Size Obtained in Independent Samples t-test of Scores on Forced Choice Triad Task

This calculated effect size of d=0.43 is, by convention, moderate (Cohen, 1992; Hojat & Xu, 2004). Using the G-Power 3.1.2 analysis program (Faul, 2009), post hoc or observed power of 0.21 was calculated for this analysis.

The pattern of responses for the two different comparison types (surface-deep and surface-deep/surface) in the forced choice triad task was analyzed for any difference between cohorts. Table 8 shows the percentage of deep matches for each type of comparison type for the two practice cohorts.

Table 8 *Deep Matches by Practice Cohort on Surface-Deep and Surface-Deep Surface Comparison Types*

	Practice Level	N (Individuals)	Mean (SD)
Percentage of Deep Matches on Comparison Type 1 (S-D)	Novice/Less Experienced	8	0.30 (0.40)
	Competent/Experienced	9	0.31 (0.23)
Percentage of Deep Matches on Comparison Type 2(S-SD)	Novice/Less Experienced	8	0.25 (0.30)
	Competent/Experienced	9	0.49 (0.30)

A two-factor mixed factorial analysis of variance (ANOVA) was conducted, with practice cohort as the between-subjects factor and comparison type as the within-subjects factor. This analysis showed that: Practice level did not have a significant effect on score [F (0.95,1=0.91), p=0.36]; comparison type did not have a significant effect on score [F (0.95, 1) =0.75, p=0.40]; nor did the interaction between practice level and comparison type have a significant effect upon score [F (0.95,1) =2.40, p=0.14]. However, the direction of the predicted difference in scores as a function of practice level and comparison type was observed. An analysis of effect size differences between the SD and S-SD comparison types across practice cohorts showed a difference between practice cohorts according to item comparison type. For the SD comparison type, an effect size across cohorts of 0.03 was calculated (Figure 5) whereas for the S-SD comparison type, an effect size across practice level of 0.8 was calculated (Figure 6).

$$\because s = \sqrt{\frac{SS_1 - SS_2}{df_1 - df_2}}$$

$$\therefore s = \sqrt{\frac{[(n_1 - 1) \times s_1^2] + [(n_2 - 1) \times s_2^2]}{df_1 - df_2}}$$

$$\therefore s = \sqrt{\frac{[(8 - 1) \times 0.4^2] + [(9 - 1) \times 0.23^2]}{8 + 9 - 2}}$$

$$= \sqrt{\frac{[7 \ \times 0.16] + \ [8 \times 0.05]}{15}}$$

$$= \sqrt{\frac{1.12 + 0.42}{15}}$$

$$= \sqrt{\frac{1.54}{15}}$$

$$= \sqrt{0.103}$$

$$= 0.32$$

$$d = \frac{\mu_1 - \mu_2}{\sigma}$$

$$= \frac{0.31 - 0.30}{0.32}$$

$$= \frac{0.01}{0.32}$$

$$= 0.03$$

Figure 5 Calculation of Effect Size for SD Comparison Type[9]

[9] μ_1 and μ_2 are approximated by the sample mean values and σ is approximated by the sample pooled standard deviation, s.

$$\because s = \sqrt{\frac{SS_1 - SS_2}{df_1 - df_2}}$$

$$\therefore s = \sqrt{\frac{[(n_1 - 1) \times s_1^2] + [(n_2 - 1) \times s_2^2]}{df_1 - df_2}}$$

$$= \sqrt{\frac{[(8-1) \times 0.30^2] + [(9-1) \times 0.30^2]}{8+9-2}}$$

$$= \sqrt{\frac{[7 \times 0.09] + [8 \times 0.09]}{15}}$$

$$= \sqrt{\frac{0.63 + 0.72}{15}}$$

$$= \sqrt{\frac{1.35}{15}}$$

$$= \sqrt{0.09}$$

$$= 0.3$$

$$\therefore d = \frac{\mu_1 - \mu_2}{\sigma}$$

$$\therefore d = \frac{0.49 - 0.25}{0.3}$$

$$= \frac{0.24}{0.3}$$

$$= 0.8$$

Figure 6 Calculation of Effect Size for S-SD Comparison Type

Using G-Power 3.1.2, a post hoc power value of 0.48 for the comparison of mean scores on the surface-surface/deep comparison type items by practice cohort was calculated.

Chapter Three

Discussion

In this chapter, the study findings, although acknowledged to be provisional, are summarized. The use of the forced choice triad task in the investigation of hypothesized perceptual differences along the novice-expert continuum is discussed. Barriers to participant recruitment are identified and recommendations to mitigate these barriers at the systems and organizational levels are made. In addition, recommendations are made for the conduct of future, multi-site studies involving nurses in acute care practice settings. Finally, areas of future research in novice-expert differences that may inform education for specialty nursing practice are discussed.

Findings

Perceptions of cardiovascular nurses at two points in the novice-expert continuum were found to differ as hypothesized in this study, although not to statistical significance. On the instrument as a whole, the mean score for alignment with the deep comparator was 0.29 for the novice/less experienced cohort and 0.41 for the competent/experienced cohort. For trials consisting of a surface-deep comparison type, the two cohorts scored similarly; for trials of the surface-surface/deep comparison type, the mean scores were 0.25 for the novice/less experienced cohort and 0.49 for the competent/experienced cohort. The latter finding failed to replicate those of Hardiman et al. (1989) and of Rabinowitz and Hogan (2002) that individuals, regardless of their position on the novice-expert continuum, tend to base similarity judgments on surface features when both referent items within a trial have surface similarities to the target item.

The small sample recruited to participate in this study precludes making definitive conclusions based on its findings. When non-significant test results are obtained with a small sample, it is possible that the null hypothesis is true; i.e., no statistically significant difference exists in scores between groups on measures of the phenomenon of interest. It is also possible that such a difference does exist in the population, but that it is too small to be detected in a small sample. A calculation of observed power has been advocated as a means of providing possible explanations for non-significant findings (Onwuegbuzie & Leech, 2004). Sun, Pan, and Wang (2011), however, argue against this approach because lower power cannot differentiate whether non-significant findings result from the lack of hypothesized differences or from sampling error.

The magnitude of the difference on overall scores, as reflected by a calculation of effect size, d, ($d=0.43$) was found to be moderate. However, the low observed power of this analysis ($1-\beta=0.21$) does not provide evidence as to whether the small sample size or the lack of a statistically significant difference in the population accounts for these findings. The difference in the mean scores in this study ($M=0.41$ for the competent/experienced cohort and $M=0.29$ for the novice/less experienced cohort) is less than that found by Devantier et al. (2008) in their study of novice-expert differences in patient categorization ($M=0.48$ for the expert cohort and $M=0.17$ for the novice cohort). Furthermore, a large effect size of $d=0.74$ was calculated for the magnitude of the difference in scores by the novice and expert cohorts in the study by Devantier et al. (2008). These comparisons suggest that a non-significant difference in scores might have been observed had a larger sample been recruited for the present study.

While the magnitude of difference between scores based on the SD comparison type items was small (d=0.03), the difference between scores on the S-SD comparison type items was large (d=0.80). A post hoc power of 0.48 for this analysis was calculated. Again, whether the difference is not significant or the whether the study had insufficient power cannot be determined.

The Use of the Forced Choice Triad Task in Novice-Expert Studies

The forced choice triad task may have limited usefulness in the study of novice-expert differences, since the method has not been used to investigate phenomena other than the hypothesized novice-expert surface-deep perceptual difference. In studies by Hardiman et al. (1989) and Rabinowitz & Hogan (2002) that used the forced choice triad task method, when both referent items within a trial had surface similarities to the target item novices and experts alike were more likely to make similarity judgments on the basis of surface features. Likewise, if no surface similarity comparison type was available, novices and experts tended to match based on deep featural similarities. Hardiman, et al. (1989) ascribed these findings to a "similarity threshold" phenomenon, in which the presence of surface similarities in both referent items favoured similarity judgments based on surface features and the absence of surface similarities favoured similarity judgments based on deep features. In the present study, this pattern of responses was not observed. A possible explanation for this is that, in the surface/surface-deep comparison type, the features designed to reflect surface similarities were *too* similar to one another and, therefore, were uniformly disregarded by the more experienced cohort. This study did not include comparison types in which only a surface or deep similarity was available and, therefore, this finding was not tested. If future researchers are to use the procedure described in this thesis, it is recommended that the study instrument be revised to include trials in which only a surface or deep comparison type is available, and that the surface/surface-deep comparison types be revised to reflect more varied surface similarities.

Scoring of responses in this study relied to a degree upon inferences made by the raters. Participants were asked to provide rationales for their responses; however, the rationales tended to be terse and/or ambiguous and, therefore, subject to inferential bias. An amendment to the research ethics application was made to allow the researcher to seek clarification of responses; however, such clarification was vulnerable to recall bias due to the delay between the participant's completion of the instrument and the clarification. If the study were conducted in person, responses could be clarified immediately, which would mitigate these biases. Future researchers using this study's instrument or another forced choice triad task are advised to develop a strategy in advance for scoring responses that do not reflect the basis on which the items were designed and those whose rationales evince faulty knowledge.

Barriers to Recruitment

Due to the snowball sampling strategy that was employed, it was not anticipated that research ethics board (REB) approval in addition to that received from McMaster University would be required by individual organizations to recruit their staff to participate in the study. The requirement to apply for individual institutional ethics approval by some organizations and not others and inconsistencies within organizations as to whether REB approval would be required resulted in considerable delay in the study timeline. In addition, when it became clear that the snowball recruitment strategy was not yielding sufficient participants, amendments to REB applications had to be submitted in order to use alternative recruitment strategies; this requirement also resulted in delays and the alternative strategies did not result in a significant increase in recruitment.

Equipoise refers to bona fide uncertainty that a clinical intervention will be of benefit; this principle underlies the conduct of ethical research (Freedman, 1987). Eva (2007) points out that the assumption that educational research is without potential harm violates the principle of equipoise. Harm may result from the opportunity cost to research participants; that is, participation in research consumes time and effort that could be expended on other activities. Harm may result from the converse of the benefits to participants that the intervention was intended to provide (Eva, 2007). A third, broader category of harm may accrue from studies lacking an adequate theoretical or empirical basis that are used to inform decisions and allocate resources (Eva, 2007). Therefore, ethics review of research intended to inform educational strategies is necessary. However, if all research ethics boards are guided by the principles of the Tri-Council Policy: Ethical Conduct of Research Involving Humans (Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, and Social Sciences and Humanities Research Council of Canada, 2010), it is uncertain as to why some organizations require ethics approval beyond the author's university ethics board to conduct research while others do not. This inconsistency is particularly puzzling when the study under consideration is a low-risk study that did not involve a treatment, personal health information, patients or vulnerable populations. If there is consensus as to the purposes of ethics oversight there should be no inconsistency in the requirement for it.

The conduct of research would be greatly facilitated by a single REB application for multi-site research. This would eliminate inconsistent requirements among organizations. Organizations could also improve communication between their REB's and nursing and other research committees in order that potential researchers do not receive conflicting information regarding ethics requirements. Organizations could also provide a uniform, expedited review process for protocol amendment applications for student research projects.

Some institutional practices likely had an effect on recruitment for participation in this study, i.e., policies that limit time for displaying posters and that require posters to be displayed in locked cabinets, so that potential participants must manually record contact information. Organizations should examine how their policies for displaying invitations to participate in research may affect recruitment and revise those policies that actually hinder it.

Varying familiarity with research conventions on the part of nurses may be ascribed to the diverse educational preparation that nurses have. The baccalaureate degree has been required for entry to practice since 2005 (Registered Nurses' Association of Ontario, n.d.); therefore, there are many nurses still in practice whose highest educational attainment is a nursing diploma. Diploma nursing programs have historically had little or no content relating to research theory or practice (Tranmer, Lochaus-Gerlach, & Lam, 2002); therefore, there are many nurses who have limited knowledge of and experience with research. Although baccalaureate programs include theoretical courses on research, they may not provide experience in the practice of research. Master's degree-prepared nurses may have completed a course-based or a thesis-based program; therefore, even among master's degree-prepared nurses there is not a uniform degree of experience with the conduct of research.

Although the reasons for nurses' low levels of participation in research can often be only inferred, it is possible that research is perceived by some nurses as an elitist activity that is irrelevant to clinical nursing practice. Rolfe, Segrott and Jordan (2008), in a study of nurses' understanding of evidence-based practice, obtained a response rate of only 8.9% to a questionnaire; this was anecdotally ascribed to nurses not having time to complete the questionnaires during working hours and being unwilling to complete them in off-working hours. Edwards et al. (2006) attribute the low response rate in their study on the influence of supervision on the levels of burnout experienced among community mental health nurses to their having been “over-researched” (Edwards et al., 2006, p.1014) and to their not perceiving the value of research.

Health care organizations generally, in their mission statements, express a valuing of the role of research in innovation and patient care. Therefore, organizational culture should be supportive of the conduct of research. If such organizational support were available, nurses might be more willing to participate in research. Tranmer, Lochaus-Gerlach, and Lam (2002) found that the use of nursing research in practice was positively correlated with organizational structures and processes that supported their participation in research. These processes and structures include education on research methods, opportunities to participate in research that is clinically relevant at all phases of the research project, protected time to participate in research, and assistance from those with research expertise (Tranmer et al., 2002). The processes and structures to promote nurses’ participation in research--as consumers, investigators, and participants--should be in place at every organization. Large, academic health science centres generally have departments of nursing research that can, and often do, support staff participation in research activities. Smaller, community organizations that do not have nursing research departments could benefit from their relationships with the local Schools of Nursing for which they provide clinical placements for nursing students; faculty could function in a consultant role to support in-house research activities. Smaller organizations could also benefit materially from the funding available through the Registered Nurses’ Association of Ontario’s

Advanced Clinical/Practice Fellowship program to support the development of nurses' knowledge of and experience with research.

The mere existence of processes and structures to support nursing participation in research is insufficient. Strong nursing leadership, both formal and informal, that recognizes the value of research to nursing practice and actively encourages the involvement of all nurses in research is necessary for the utilization of supportive processes and structures.

The reform of some of the systemic and organizational policies regarding the conduct of research is clearly needed. However, such reform may not be imminent and, therefore, current researchers must contend with the status quo. As noted, a major barrier to participant recruitment in this study was the inconsistency in the requirement for institutional research ethics approval, both among and within organizations. Other researchers who wish to conduct studies similar to the reported study might consider applying for REB approval early in their study timeline from all organizations from which they intend to invite staff participation, regardless of the advisement of its necessity. They might also consider using a variety of recruitment strategies and tracking their success; this would allow researchers to focus their efforts on the successful strategies early in the recruitment phase. Researchers could explore how recruitment might be done via professional associations: The Canadian Council of Cardiovascular Nurses publishes a journal that all members receive in which the study could be advertised. Other potential researchers also might contemplate using different methods of administering the task; i.e., in person with a group, online, and by mail. If resources permit, researchers could also offer larger incentives to make participation more attractive.

Researching Novice-Expert Differences in Nursing Practice

The findings of the present study, although they are acknowledged to be provisional, do not generally refute the findings of similar studies. Novice-expert studies are conducted partly in the hope that knowledge about such differences will permit valid inferences about how individuals progress toward expertise in a domain; such inferences, it is hoped, can inform strategies to assist individuals to advance along this trajectory. Much prior research into novice-expert differences in nursing practice have relied upon criteria for practice level and methods that have been challenged. Future research must acknowledge these criticisms. If, as Ericsson insists (2006a), the only valid criterion by which practice level may be classified is consistent performance on domain-relevant tasks, the researcher's aim would be to first identify such a task, to specify criteria that characterize performance of that task at different levels of practice, to identify those who meet these criteria and analyze how they differ from those who do not meet the criteria. Within cardiovascular nursing, such a domain-relevant task might be hemodynamic monitoring and the use of data derived from it and the integration of that data with other data such as laboratory values and physical assessment findings to guide practice.

The preferred method for evaluating performance on domain-relevant tasks might be a think-aloud protocol as advocated by Ericsson (2006b). A think-aloud protocol could be conducted in an authentic clinical setting or in a simulated setting; both have limitations. In authentic clinical settings, participants' spontaneous verbalizations might be constrained by the presence of conscious patients or their family members. Post hoc reporting of thought processes would be subject to recall bias. A think-aloud protocol conducted in a simulated environment would eliminate such concerns; however, regardless of the fidelity of the simulation, it is questionable that a scenario as complex and highly nuanced as a real clinical situation could be created. As Mylopoulos and Regehr (2011) point out, cognition and learning are influenced by the context within which they occur; to study these phenomena out of their natural context ignores the situated character of cognition and learning. A trajectory along which the novice progresses toward expertise is inferred from the study of expert performance within a domain (Lajoie, 2003) or from comparisons of novice and expert performance (Daley, 1999). Longitudinal studies that track the professional development of individuals would make this trajectory explicit and would contribute to an understanding of the contextual factors that support or hinder the development of expert practice. Such contextual factors include the access to learning opportunities both within the clinical area and in continuing education programs offered in the community, the influence of mentors, the degree of commitment by the nursing leadership to staff development, and the influence of the unit culture upon professional development. It is plausible that these factors affect the degree of motivation that nurses have to engage in the sustained, deliberate practice that Ericsson (2006a) insists is necessary for the development of expertise.

Novice-expert differences in cognition and performance in a variety of domains have been ascribed to both the quantity and the nature of the organization of domain knowledge in memory. Regehr and Norman (1996) propose that the central challenge of devising educational strategies entails enhancing memory of domain-specific knowledge; they suggest that this objective might be accomplished by making information meaningful, by minimizing context-dependence of learning, and by providing opportunities to retrieve domain knowledge frequently. Mayer (2010) states that meaningful learning occurs when the learner selects words and images to attend to, organizes words and images in working memory, and integrates the organization of words and images with one another and with prior knowledge in long-term memory. Meaningful learning, then, is an active process of schemata construction. Dufresne, Gerace, Hardiman, & Mestre (1992) contend that schemata, or the organization of domain-specific knowledge in memory, allow experts to make a correct judgment as to the principles that apply to a problem solution and to select the correct procedures that instantiate those principles. They conducted a series of experiments on the effect of a decision analysis tool on novices' judgments of solution similarity and problem solving; the tool prompted the novices to first identify the principle required to solve a physics problem and then to select the procedures required to solve the problem. They found that this guidance was more effective than self-directed efforts for problem-solving and schema acquisition. Goldstone (2006) proposes that "perceptual priming" by drawing explicit attention to domain principles supports transfer of learning by providing cues that are lacking in abstract representations of domain principles, such as equations. Such priming, he argues, influences how a situation is perceived and, therefore, what information is retrieved in the interpretation of the situation. Mestre (2002) suggests that when students are required to generate an abstraction to describe a phenomenon, although these abstractions are usually faulty, they are "primed" to appreciate the correct abstraction when it is presented. Hoffman, Feltovich, Fiore, Klein, and Ziebell (2009) argue against a "blocked" type of practice or a stagewise progression in instruction from simple to complex and against separating theory from practice; they contend that learning more complex principles initially may be more difficult but results in greater understanding and transfer ultimately.

Shackell and Gillespie (2009) developed an Oxygen Supply and Demand Framework for use in critical care nursing curricula. This framework was designed to make broad physiologic concepts and the links between them explicit; learners are encouraged to attempt to account for their clinical assessment findings based on these concepts. At the time of writing, Shackell and Gillespie had reported only anecdotal evidence supporting the educational efficacy of their framework (2009). Future empirical research might compare specialty nursing educational outcomes for students prepared using traditional, "stepwise" curricula and those prepared using integrative frameworks consisting of domain principles or concepts. Challenges inherent in such research are the identification of outcomes to compare and of measures to compare them.

Conclusion

In the present study, the perceptions of cardiovascular nurses at two points in the novice-expert continuum were found to differ as hypothesized, although not to statistical significance. When presented with clinical scenarios, those nearer the expert end of the continuum generally perceived deep, relational, or structural features to a greater degree than did those nearer the novice end of the continuum. Future research on novice-expert differences in nursing must reflect recognition of the criticisms made of the assumptions and methods used in previous research. The implications of the findings of research on novice-expert differences in a variety of domains for the development of educational strategies for specialty nursing practice provide rich opportunity for future research.

References

Aitken, L. (2000). Expert critical care nurses' use of pulmonary artery pressure monitoring. *Intensive and Critical Care Nursing, 16*(4), 209-220. doi: 10.1054/iccn.2000.1498

Allen, D., & Cloyes, K. (2005). The language of 'experience' in nursing research. *Nursing Inquiry*, *12*(2), 98-105. doi: 10.1111/j.1440-1800.2005.00259

Altmann, T. (2007). An evaluation of the seminal work of Patricia Benner: Theory or philosophy? *Contemporary Nurse, 25*(1-2), 114-123. doi: 10.5172/conu.2007.25.1-2.114

Benner, P. (2001). From novice to expert*: Excellence and power in clinical nursing practice. Commemorative edition.* Upper Saddle River, NJ: Prentice Hall.

Billings, D., Jeffries, P., Daniels, D., Rowles, C., Stone, C., & Stephenson, E. (2006). Developing and using online courses to prepare nurses for employment in critical care. *Journal for Nurses in Staff Development, 22*(2), 87-92. doi:10.1097/00124645-200603000-00008

Bobay, K. (2004). Does experience really matter? *Nursing Science Quarterly, 17*(4), 313-316. doi: 10.1177/0894318404269369

Bourgault, A., & Smith, S. (2004). The development of multi-level critical care competency statements for self-assessment by ICU nurses. *Dynamics*, *15*(4), 15-18.

Browner, W.S., Newman, T. B., & Hulley, S. B. (2007). Estimating sample size and power: Applications and examples. In S. Hulley, S. Cummings, W. Browner, D. Grady, & T. Newman (Eds.), *Designing clinical research* (pp.65-94). Philadelphia, PA: Lippincott, Williams, & Wilkins.

Canadian Council of Cardiovascular Nurses (November, 2008). *Standards for cardiovascular nursing.* Retrieved October 9, 2009 from http://www.cccn.ca/eng/documents/ 2009_standards_cardiovascular_nursing.pdf

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada,and Social Sciences and Humanities Research Council of Canada, Tri-Council (2010). *Policy statement: Ethical conduct for research involving humans.* Retrieved from Government of Canada website: http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/Default/

Carlson, L., Crawford, N., & Contrades, S. (1989). Nursing student novice to expert: Benner's research applied to education. *Journal of Nursing Education, 28*(4), 188-190.

Carmines, E., & Zellers, R. (1979). *Reliability and validity assessment.* (Sage University Paper Series on Quantitative Applications in the Social Sciences, series n. 07-017). Thousand Oaks, CA: Sage.

Cash, K. (1995). Benner and expertise in nursing: A critique. *International Journal of Nursing Studies*, *32*(6), 527-534. doi: 10.1016/0020-7489(95)00011-3

Cavanaugh, D., & Huse, A. (2004). Surviving the nursing shortage: Developing a nursing orientation program to prepare and retain intensive care unit nurses. *Journal of Continuing Education in Nursing, 35*(6), 251-256. Retrieved from http://search.proquest.com.libaccess.lib.mcmaster.ca/docview/223320630

Chesnutt, B., & Everhart, B. (2007). Meeting the needs of graduate nurses in critical care orientation: Staged orientation program in surgical intensive care. *Critical Care Nurse, 27*(3). 36-40. Retrieved from http://ccn.aacnjournals.org.libaccess.lib.mcmaster.ca/content/27/3.toc

Chi, M.T. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P.J. Feltovich & R.R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp.21-30). New York: Cambridge University Press.

Chi, M., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science,5* (2), 121-152. doi: 10.1207/s15516709cog0502_2

Clarke, S., & Aiken, L. (2003). Failure to rescue. *American Journal of Nursing, 103*(1). 42-47. doi:10.1111/j.1440-1800.2004.00210

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. doi:10.1037//0033-2909.112.1.155

College of Nurses of Ontario (2008). Authorizing mechanisms. Compendium of Standards of Practice for Nurses in Ontario. [Electronic version]. Retrieved from http://www.cno.org/Global/docs/prac/41075_AuthorizingMech.pdf

Collins, M., & Thomas, T. (2005). Creation of a stepdown nurse internship program. *Journal for Nurses in Staff Development, 21*(3), 115-119. doi:10.1097/00124645-200505000-00007

Cook, N. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies, 41*(6), 801-849. doi: 10.1006/ijhc.1994.1083

Corcoran-Perry, S., Narayan, S., & Cochrane, S. (1999). Coronary care nurses' clinical decision making. *Nursing and Health Sciences, 1*(1), 49-61. doi: 10.1046/j.1442-2018.1999.00008.x

Cote, D., & Burwell, K. (2007). A revised nephrology nurses' clinical ladder. *Nephrology Nursing Journal, 34*(2), 243-248. Retrieved from http://search.proquest.com. libaccess.lib .mcmaster.ca/ docview/216537273 /fulltextPDF/

Crandall, B., & Getchell-Reiter, K. (1993). Critical decision method: A technique for elicting concrete assessment indicators from the inuition of NICU nurses. *Advances in Nursing Science, 16*(1), 42-51. doi:10.1145/63266.63288

Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302. doi 10.1037/h0040957

Currey, J., & Botti, M. (2006). The influence of patient complexity and nurses' experience on haemodynamic decision-making following cardiac surgery. *Intensive and Critical Care Nursing, 22*(4), 194-205. doi: 10.1016/j.iccn.2005.06.005

Custers, E., Regehr, G., & Norman, G. (1996). Mental representations of medical diagnostic knowledge: A review. *Academic Medicine, 71*(10), S55-S61. doi:10.1097/00001888-199610000-00044

Daley, B. (1999). Novice to expert: An exploration of how professionals learns. *Adult Education Quarterly, 49*(4), 133-147. doi: 10.1177/074171369904900401

Dall'Alba, G., & Sandberg, J. (2006). Unveiling professional development: A critical review of stage models. *Review of Educational Research, 76*(3), 383-412. doi: 10.3102/00346543076003383

Daniel, W.W. (2005). Estimation. In *Biostatistics: A foundation for analysis in the health sciences.*(pp.156-210). Hoboken, NJ: John Wiley & Sons.

Devantier, S., Minda, J., Hadarra, W., & Goldszmidt, M. (2008). Physician`s use of deep features: Expertise differences in patient categorization. In B.C. Love, K. McRae, & V.M. Soutsky (Eds.), Proceedings of the 30th Annual Conference of the Cognitive Science Society (pp. 1687-1692). Austin, TX: Cognitive Science Society

Dillman, D. A.(2007). *Mail and internet surveys: The tailored design method,* 2nd Ed. New York: Wiley.

Domain (n.d.). In *Merriam-Webster's online dictionary.* Retrieved from http://www.merriam-webster.com/dictionary/domain

Dreyfus, S. (2004). The five-stage model of adult skill acquisition. *Bulletin of Science, Technology & Society*, *24*(3), 177-181. doi: 10.1177/0270467604264992

Dreyfus, H., & Dreyfus, S. (1984). *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: Free Press.

Dufresne, R., Gerace, W., Hardiman, P., & Mestre, J. (1992). Constraining novices to perform expertlike problem analyses: Effects on schema acquisition. *The Journal of the Learning Sciences, 2*(3), 307-331. doi:10.1207/s15327809jls0203_3

Ebright, P., Patterson, E., Chalko, B., & Render, M. (2003). Understanding the complexity of registered nurse work in acute care settings. *Journal of Nursing Administration, 33*(12), 630-638. doi:10.1097/00005110-200312000-00004

Edwards, D., Burnard, P., Hannigan, B., Cooper, L., Adams, J., Juggessur, T., Fothergil, A., & Coyle, D. (2006). Clinical supervision and burnout: the influence of clinical supervision for community health nurses. *Journal of Clinical Nursing, 15*(8), 1007-1015. doi: 10.1111/j.1365-2702.2006.01370

English, I. (1993). Intuition as a function of the expert nurse: A critique of Benner's novice to expert model. *Journal of Advanced Nursing, 18*(3), 387-393. doi: 10.1046/j.1365-2648.1993.18030387

Ericsson, A. (2006a). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P.J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 683-703). New York: Cambridge University Press.

Ericsson, A. (2006b). Protocol analysis and expert thought. In K. A. Ericsson, N. Charness, P.J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 223-241). New York: Cambridge University Press.

Ericsson, K., & Simon, H. (1980). Verbal reports as data. *Psychological Review, 87* (3), 215-251.doi 10.1037/0033-295X.87.3.215

Ericsson, K., Whyte, J., & Ward, P. (2007). Expert performance in nursing: Reviewing research on expertise in nursing within the framework of the expert-performance approach. *Advances in Nursing Science, 30*(1), E58-E71. Retrieved from http://ovidsp.tx.ovid.com.libaccess.lib.mcmaster.ca/sp-3.4.2a/ovidweb.cgi?

Eva, K. (2007). The yin and yang of educational research. *Medical Education, 41*(8), 724-725, doi 10.1111/j.1365-2923.2007.02806.x

Eva, K.W., Brooks, L.R, Norman, G.R. (2002). Forward reasoning as a hallmark of expertise in medicine: Logical, psychological, and phenomenological inconsistencies. In S.Shohov (Series Ed.), *Advances in Psychology Research Series: Vol. 8*, (pp.41-69). New York : Nova Science Publishers, Inc.

Eva, K., Neville, A., & Norman, G. (1998). Exploring the etiology of content specificity: Factors influencing analogical transfer and problem-solving. *Academic Medicine*, *73*(10), S1-S5. doi:10.1097/00001888-199810000-00028

Everson, F. (2005). Regional emergency nursing program: How one region solved its need for more efficient orientation education. *Journal of Emergency Nursing*, *31*(4), 398-399. doi: 10.1016/j.jen.2005.06.002

Falan, S.L. (2007). *Similarity discernnment in nursing and general representations*. (Doctoral dissertation). Retrieved from http://proquest.umi.com.libaccess.lib.mcmaster.ca/

Falan, S. (2010). Concept analysis of similarity applied to nursing diagnoses: Implications for educators. *International Journal of Nursing Terminologies and Classifications*, *21*(4), 144-155. doi: 10.1111/j.1744-618X.2010.01163.x

Farrington-Darby, T., & Wilson, J. (2006). The nature of expertise: A review. *Applied Ergonomics*, *37*(1), 17-32. doi:10.1016/j.apergo.2005.09.001

Faul, F. (2009). GPower (Version 3.1.2). [Computer software]. Universitat Kiel.

Feldon, D. (2007). The implications of research on expertise for curriculum and pedagogy. *Educational Psychology Review*, *19*(2), 91-110. doi:10.1007/s10648-006-9009-0

Freedman, B. (1987). Equipoise and the ethics of clinical research. *New England Journal of Medicine*, *317*(3), 141-5. doi 10.1056/NEJM198707163170304

Gatley, E. (1992). PREPP: From novice to expert; post-registration education and practice report. *British Journal of Nursing, 1*(2), 88-91. doi:10.1016/0260-6917(92)90032-J

Gavlak, S. (2007). Centralized orientation: Retaining graduate nurses. *Journal for Nurses in Staff Development, 23*(1), 26-30. doi:10.1097/00124645-200701000-00006

Gentner, D. (1988). Metaphor as structure-mapping: The relational shift. *Child Development, 59* (1), 47-59. doi:10.1111/j.1467-8624.1988.tb03194.x

Gentner, D., & Markman, A. (1994). Structual alignment in comparison: No difference without similarity. *Psychological Science, 5*(3), 152-158. doi:10.1111/j.1467-9280.1994.tb00652.x

Gentner, D., Rattermann, M., & Forbus, K. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology, 25*(4), 524-575. doi:10.1006/cogp.1993.1013

Gobet, F., & Chassy, P. (2008). Towards an alternative to Benner's theory of expert intuition in nursing: A discussion paper. *International Journal of Nursing Studies, 45*(1), 129-139. doi: 10.1016/j.ijnurstu.2007.01.005

Goldstone, R. (2006). The complex systems see-change in education. *The Journal of the Learning Sciences, 15*(1), 35-43. doi:10.1207/s15327809jls1501_5

Goldstone, R., & Son, J. (2005). Similarity. In K.J. Holyoak & R.G. Morrison (Eds.), Cambridge handbook of thinking and reasoning (pp. 13-36). New York, NY: Cambridge University Press.

Graham, P. (2006). Experience with a community-based consortium approach. *Critical Care Nursing Quarterly, 29*(3), 207-217. Retrieved from http://ovidsp.tx.ovid.com.libaccess. lib.mcmaster. ca/sp-3.4.2a/ovidweb.cgi?QS2=434f4e1a

Hardiman, P., Dufresne, R., & Mestre, J. (1989). The relation between problem categorization and problem solving among experts and novices. *Memory and Cognition, 15*(5), 627-638. doi:10.3758/BF03197085

Harrelson, M. (2007). An innovative approach to orient nursing students. *Journal for Nurses in Staff Development, 23*(2). doi:10.1097/01.NND.0000266612.16434.4c

Health Canada (2006). *Canada's health care system*. Retrieved from http://www.hc-sc.gc.ca/hcs-sss/pubs/system-regime/2005-hcs-sss/index-eng.php

Henderson-Everhardus, M.C. (2004). *Does nursing experience contribute to the accuracy of vascular assessment in the detection of peripheral arterial disease?* (Master's thesis). Retrieved from http://proquest.umi.com.libaccess.lib.mcmaster.ca

Henning, S., & Cohen, E. (2008). The competency continuum: Expanding the case manager's skill sets and capabilities. *Professional Case Management, 13*(3), 127-148. doi: 10.1097/01.PCAMA.0000319966.30864.bf

Hmelo-Silver, C., Marathe, S., & Liu, L. (2007). Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. *The Journal of the Learning Sciences, 16*(3), 307-331. doi: 10.1080/10508400701413401

Hmelo-Silver, C., & Pfeffer, M. (2004). Comparing expert and novice understanding of a complex system from the perspective of structure, behaviours, and functions. *Cognitive Science, 28*(1). 127-138. Retrieved from http://resolver.scholarsportal.info.libaccess.lib. mcmaster.ca/ resolve/03640213/v28i0001/127_ceanuoposbaf

Hoffman, R., Feltovich, P., Fiore, S., Klein, G., & Ziebell, D. (2009). Accelerated learning? *IEEE Intelligent Systems, 24*(2), 18-22. doi 10.1109/MIS.2009.21

Hoffman, R.R., & Lintern, G. (2006) Eliciting and representing the knowledge of experts. In K. Ericsson, N. Charness, P. Feltovich, & R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp.203-222). New York: Cambridge University Press.

Hoffman, R., Shadbolt, N., Burton, A., & Klein, G. (1995). Eliciting knowledge from experts: A methodological analysis. *Organizational Behaviour and Human Decision Processes, 62*(2), 129-158. doi:10.1006/obhd.1995.1039

Hojat, M., & Xu, G. (2004). A visitor's guide to effect sizes: Statistical significance versus practical (clinical) importance of research findings. *Advances in Health Sciences Education, 9*(3), 241-249. doi:10.1023/B:AHSE.0000038173.00909.f6

Holyoak, K., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition, 15*(4), 332-340. doi:10.3758/BF03197035

Horgan, D., Millis, K., & Neimeyer, R. (1989). Cognitive reorganization and the development of chess expertise. *International Journal of Personal Construct Psychology, 2*(1), 15-36. doi:10.1080/08936038908404735

Jacobson, A., & Winslow, E. (2005) Variables influencing intravenous catheter insertion difficulty and failure: An analysis of 339 intravenous catheter insertions. *Heart & Lung, 34*(5), 345-359. doi: 10.1016/j.hrtlng.2005.04.002

Johnson, K., & Mervis, C. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology, 126*(3), 248-277. doi: 10.1037/0096-3445.126.3.248

Kinnear, P.P., & Gray, C. D. (2009). *SPSS 16 made simple*. New York, NY: Psychology Press.

Kirk, E., & Ashcraft, M. (2001). Telling stories: The peril and promise of using verbal reports to study math strategies. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27*(1), 157-175. doi: 10.1037/0278-7393.27.1.157

Kutney-Lee, A., Lake, E., & Aiken, L. (2009). Development of the hospital nurse surveillance capacity profile. *Research in Nursing and Health, 32*(2), 217-228. doi: 10.1002/nur.20316

Lajoie, S. (2003). Transitions and trajectories for studies of expertise. *Educational Researcher, 32*(8), 21-25. doi: 10.3102/0013189X032008021

Larew, C., Lessans, C., Spunt, D., Foster, D., & Covington, B. (2006). Innovations in clinical simulation: Application of Benner's theory in an interactive patient care simulation. *Nursing Education Perspectives, 27*(1), 16-21. Retrieved from http://go.galegroup.com.libaccess.lib.mcmaster.ca/ps/i.do?id=GALE%7CA142786652&v=2.1&u=ocul_mcmaster&it=r&p=AONE&sw=w

Lavoie-Trembley, M., Viens, C., Forcier, M., Labrosse, N., Lafrance, M., Laliberte, D., & Lebeuf, M. (2002). How to facilitate the orientation of new nurses into the workplace. *Journal for Nurses in Staff Development, 18*(2), 80-85. doi:10.1097/00124645-200203000-00006

Lenthe, R. (2001). Some practical guidelines for effective sample size calculations. *The American Statistician, 55*(3), 187-193. doi:10.1198/000313001317098149.

Lincoln, Y., & Guba, E. (1985). *Naturalistic enquiry*. Newbury Park, Ca: Sage.

Loiseau, D., Kitchen, K., & Edgar, L. (2003). A comprehensive ED orientation for new graduates in the emergency department: The four-year experience of one Canadian teaching hospital. *Journal of Emergency Nursing, 29*(6), 522-527. doi:10.1016/j.jen.2003.09.010

Lott, T. (2006). Creating a new nursing services orientation program. *Journal for Nurses in Staff Development, 22*(5), 214-221. Retrieved from http://ovidsp.tx.ovid.com. libaccess. lib.mcmaster. ca/sp-3.4.2a/ovidweb.cgi?QS2=434f4e1a73d37

Lyneham, J., Parkinson, C., & Denholm, C. (2008). Explicating Benner's concept of expert practice: Intuition in emergency nursing. *Journal of Advanced Nursing, 64*(4), 380-387. doi: 10.1111/j.1365-2648.2008.04799

Lynn, M. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382-386. doi:10.1097/00006199-198611000-00017

Manuel, D. G., Leung, M., Nyugen, K., Tanuseputro, P. &. Johansen, H. (2006). The burden of cardiovascular disease in Canada. In J.V. Tu, W.A. Ghali, L.Pilote , & S. Brien (Eds.), *Canadian Cardiovascular Atlas* (pp. 15-22). Toronto, ON: Pulsus Group.

Marcum, D., & West, R. (2004). Structured orientation for new graduates: A retention strategy. *Journal for Nurses in Staff Development, 20*(3), 118-124. doi:10.1097/00124645-200405000-00003

Mayer, R. (2010). Applying the science of learning to medical education. *Medical Education*, *44*(6), 543-549. doi: 10.1111/j.1365-2923.2010.03624.x

McClure, J., Sonak, B., & Suen, H. (1999). Concept map assessment of classroom learning: Reliability, validity and logistical practicality. *Journal of Research in Science Teaching, 36*(4), 475-492. doi: 10.1002/(SICI)1098-2736(199904)36:4<475::AID-TEA5>3.0.CO;2-O

McGaghie, W., McCrimmon, D., Mitchell, G., & Thompson, J. (2004). Concept mapping in pulmonary physiology using Pathfinder scaling. *Advances in Health Sciences Education, 9*, 225-240. doi: 10.1023/B:AHSE.0000038299.79574.e8

McKane, C. (2004). Clinical objectives: A method to evaluate clinical performance in critical care orientation. *Journal for Nurses in Staff Development, 20*(3), 134-139. doi:10.1097/00124645-200405000-00006

Medin, D. (1989). Concepts and conceptual structure. *American Psychologist, 44*(12), 1469-1481. doi:10.1037//0003-066X.44.12.1469

Medin, D., Goldstone, R., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, *100*(2), 254-278. doi:10.1037//0033-295X.100.2.254

Mestre, J. (2002, March). Transfer of learning: Issues and research agenda. Report of a workshop held at the National Science Foundation, Arlington, VA. Retrieved from http://www.nsf.gov/pubs/2003/nsf03212/nsf03212.pdf

Meyer, T., & Xu, Y. (2005). Academic and clinical dissonance in nursing education: Are we guilty of failure to rescue? *Nurse Educator, 30*(2), 76-79. doi:10.1097/00006223-200503000-00010

Mylopoulos, M., & Regehr, G. (2011). Putting the expert together again. *Medical Education, 45*(9), 920-926. doi 10.1111/j.1365-2923.2011.04032.x

Nelson, S., & McGillion, M. (2004). Expertise or performance? Questioning the rhetoric of contemporary narrative use in nursing. *Journal of Advanced Nursing,47*(6), 631-638. doi: 10.1111/j.1365-2648.2004.03151

Nolan, S., & Murphy, J. (2006). The tried , true, and new: A traditional approach to critical care education with a new twist. *Critical Care Nursing Quarterly*, *29*(3), 199-206. Retrieved from http://ovidsp.tx.ovid.com.libaccess.lib.mcmaster.ca/sp3.4.2a/ovidweb.cgi?QS2=434

Norman, G. (2005). Research in clinical reasoning: Past history and current trends. *Medical Education, 39*(4), 418-427. doi:10.1111/j.1365-2929.2005.02127.x

Norman, G. (2006). Building on experience: The development of clinical reasoning. *New England Journal of Medicine, 355*(21), 2251-2252). doi:10.1056/NEJMe068134

Norman, G. (2008). Problem-based learning makes a difference. But why? *Canadian Medical Association Journal, 178*(1), 61-62. doi:10.1503/cmaj.071590

Norman, G.R., & Streiner, D.L. (2008). The nature of data and statistics. In *Biostatistics: The bare essentials*. Hamilton, ON: B.C. Decker.

Nurse Practitioners' Association of Ontario (n.d.). *History of NP Role Development in Ontario.* Retrieved April 25, 2009 from http://www.npao.org/history. aspx.

Olson, J., & Rueter, H. (1987) Extracting expertise from experts: Methods for knowledge acquisition. *Expert Systems, 4*(3), 152-168). doi:10.1111/j.1468-0394.1987.tb00139.x

Onwuegbuzie, A., & Leech, N. (2004). Post hoc power: A concept whose time has come. *Understanding Statistics, 3*(4), 201-230. doi:10.1207/s15328031us0304_1

Paley, J. (1996). Intuition and expertise: Comments on the Benner debate. *Journal of Advanced Nursing, 23*(4), 665-671. doi:10.1111/j.1365-2648.1996.tb00035.x

Patel, V., Arocha, J., & Kaufman, D. (2001). A primer on aspects of cognition for medical informatics. *Journal of the American Medical Informatics Association, 8*(4), 324-343. doi:10.1136/jamia.2001.0080324

Patel, V., Glaser, R., Arocha, J. (2000). Cognition and expertise: Acquisition of medical competence. *Clinical and Investigative Medicine, 23*(4), 256-250. Retrieved from http://libaccess.mcmaster.ca.libaccess.lib.mcmaster.ca/login?url=http://search.proquest.com.libaccess.lib.mcmaster.ca/docview/196420776?accountid=12347

Patton, M. (2002). Qualitative interviewing. In M.Q. Patton (Ed.), *Qualitative research and evaluation methods* (pp.339-348). Thousand Oaks, CA: Sage.

Penprase, B. (2000). Collaboratively developing an orientation program for OR nurses. *Association of Operating Room Nurses Journal*, *72*(4), 663-670. doi:10.1016/S0001-2092(06)61233-7

Peterson, K., & Van Buren, K. (2006). Implementing essentials of critical care orientation: One hospital's experience with an online critical care course. *Critical Care Nursing Quarterly, 29*(3), 218-230. Retrieved from http://ovidsp.tx.ovid.com.libaccess.lib.mcmaster.ca/sp-3.4.2a/ovidweb.cgi?QS2=434f4e1a73d37e8ceb49fbbfe17c3e8f7da1e5fc6803a5eff09e35f

Pine, R., & Tart, K. (2007). Return on investment: Benefits and challenges of a baccalaureate nurse residency program. *Nursing Economics, 25*(1), 13-39. Retrieved from http://go.galegroup.com.libaccess.lib.mcmaster.ca/ps/publicationSearch.do?lm=&inPS=tr

Polit, D.F., Beck, C., & Hungler, B.P. (2001). *Essentials of nursing research: Methods, appraisal, and utilization.* Philadelphia, PS: Lippincott.

Pope, C., & Mays, N. (1995). Reaching the parts other methods cannot reach: An introduction to qualitative methods in health and health services research. *British Medical Journal, 310*(6996), 42-49.

Proulx, D., & Bourcier, B. (2008). Graduate nurses in the intensive care unit: An orientation model. *Critical Care Nurse, 28*(4), 44-52. Retrieved from http://ccn.aacnjournals.org.libaccess.lib.mcmaster.ca/content/28/4/44.full.pdf+html

Quilici, J., & Mayer, R. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology, 88*(1), 144-161. doi:10.1037//0022-0663.88.1.144

Rabinowitz, M., & Hogan, T. (2002). Using a triad judgment task to examine the effect of experience on problem representation in statistics. In J.D. Moore, K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 826-830). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Rabinowitz, M., & Hogan., T. (2009). Teacher expertise and the development of a problem representation. *Educational Psychology, 29*(2), 153-169. doi:10.1080/01443410802613301

Ragsdale, M., & Mueller, J. (2005). Plan, do, study, act model to improve an orientation program. *Journal of Nursing Care Quality, 20*(3), 268-272. doi:10.1097/00001786-200507000-00013

Rauen, C. (2004). Simulation as a teaching strategy for nursing education and orientation in cardiac surgery. *Critical Care Nurse, 24*(3), 46-51. Retrieved from http://ccn.aacnjournals.org.libaccess.lib.mcmaster.ca/content/24/3/46.full.pdf+html

Regehr, B., & Norman, G. (1996). Issues in cognitive psychology: Implications for professional education. *Academic Medicine, 71*(9), 988 -1001. doi:10.1097/00001888-199609000-00015

Registered Nurses' Association of Ontario (n.d.). *The Degree Requirement.* Retrieved from http://www.rnao.org/Page.asp?PageID=924&ContentID=1656

Reischman, R., & Yarandi, H. (2002). Critical care cardiovascular nurse expert and novice diagnostic cue utilization. *Journal of Advanced Nursing, 39*(1), 24-34. doi: 10.1046/j.1365-2648.2000.02239

Roach, K. (2006). Measurement of health outcomes: Reliability, validity, and responsiveness. *Journal of Prosthetics and Orthotics,18*(1), 8-12. Retrieved from http://www.oandp.org/jpo/library/2006_01S_008.asp

Rolfe, G. (1997). Science, abduction and the fuzzy nurse: An exploration of expertise. *Journal of Advanced Nursing, 25*(5), 1070-1075. doi: 10.1046/j.1365-2648.1997.19970251070

Rolfe, G., Segrott, J., & Jordan, S. (2008). Tensions and contradictions in nurses' perspectives of evidence-based practice. *Journal of Nursing Management*, *16*(4) 440-451. doi: 10.1111/j.1365-2834.2008.00864

Rudge, T. (1992). Reflections on Benner: A critical perspective. *Contemporary Nurse, 1*(2), 84-88. doi: 10.5172/conu.1.2.84

Ruiz-Primo, M., & Shavelson, R. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching, 33*(6), 569-600. doi: 10.1002/(SICI)1098-2736(199608)33:6<569::AID-TEA1>3.0.CO;2-M

Santucci, J. (2004). Facilitating the transition into nursing practice: Concepts and strategies for mentoring new graduates. *Journal for Nurses in Staff Development, 20*(6), 274-284. doi:10.1097/00124645-200411000-00007

Schmidt, H. (2004). Alternative approaches to concept mapping and implications for medical education: Commentary on reliability, validity and future research directions. *Advances in Health Sciences Education, 9*(3), 251-256. doi: 10.1023/B:AHSE.0000038309.92212.44

Schmidt, H., Norman, G., & Boshuizen, H. (1990). A cognitive perspective on medical expertise: Theory and implications. *Academic Medicine, 65*(10), 611-621. doi:10.1097/00001888-199010000-00001

Shackell, E., & Gillespie, M. (2009). The oxygen supply and demand framework: A tool to support integrative learning. *Dynamics, 20*(4), 15-19. Retrieved from http://go.galegroup.com.libaccess.lib.mcmaster.ca/ps/i.do?&id=GALE%7CA213694477&v=2.1&u=ocul_mcmaster&it=r&p=AONE&sw=w

Shapiro, M. (1998). A career ladder based on Benner's model: An analysis of expected outcomes. *Journal of Nursing Administration, 28*(3), 13-19. doi:10.1097/00005110-199803000-00005

Shirley, M. (2007). Competencies and tips for effective leadership: From novice to expert. *Journal of Nursing Administration, 37*(4), 167-170. doi: 10.1097/01.NNA.0000266842.54308.38

Sloman, S., & Rips, L. (1998). Similarity as an explanatory construct. *Cognition, 65*(2), 87-101. doi:10.1016/S0010-0277(97)00048-6

Square, N. (2010). Modeling clinical applications in intensive care settings for nursing orientation. *Advances in Neonatal Care, 10*(6), 325-329. doi: 10.1097/ANC.0b013e3181fc881b

Squires, A. (2002). New graduate orientation in the rural community hospital. *The Journal of Continuing Education in Nursing, 33*(5), 203-209. Retrieved from http://libaccess.mcmaster.ca.libaccess.lib.mcmaster.ca/login?url=http://search.proquest.com.libaccess.lib.mcmaster.ca/docview/223324411?accountid=12347

Storey, S., Crist, L, Nelis, D., Murphy, L., & Fisher, M. (2007). Creation of a career enhancement program for a hospital-based education and development department. *Journal for Nurses in Staff Development, 23*(2), 55-61. doi:10.1097/01.NND.0000266610.70691.3b

Sun, S., Pan, W., & Wang, L. (2011). Rethinking observed power: Concept, practice, and implications. *Methodology 7*(3), 81-87. doi 10.1027/1614-2241/a000025

Tabak, N., Bar-Tal, Y., & Cohen-Mansfield, J. (1996) Clinical decision making of experienced and novice nurses. *Western Journal of Nursing Research, 18*(5), 534-547. doi: 10.1177/019394599601800505

Taylor, C. (2002). Assessing patient's needs: Does the same information guide expert and novice nurses? *International Nursing Review, 49*, 11-19. doi: 10.1046/j.1466-7657.2002.00098

Taylor, E., & Hummel, J. (2009). Finding similarity in a model of relational reasoning. *Cognitive Systems Research*, *10*(3), 229-239. doi:10.1016/j.cogsys.2008.09.004

Tolhurst, B., & Bonner, A. (2000). Development of clinical assessment criteria for postgraduate nursing students. *Collegian: Journal of the Royal College of Nursing, Australia*, *7*(2), 20-25. doi:10.1016/S1322-7696(08)60361-3

Tranmer, J., Lochaus-Gerlach, J., & Lam, M. (2002). The effect of staff nurse participation in a clinical nursing research project on attitude towards, access to, support of and use of research in the acute care setting. *Canadian Journal of Nursing Leadership, 15*(1), 18-26. Retrieved from http://www.longwoods.com.libaccess.lib.mcmaster.ca/content/19137

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*(4), 327-352. doi: 10.1037/0033-295X.84.4.327

Ung, L., Cook, S., Edwards, B., Hocking, L., Osmond, F., & Buttergieg, H. (2002). Peripheral intravevous cannulation in nursing: Performance predictors. *Journal of Infusion Nursing, 25*(3). 189-195. doi:10.1097/00129804-200205000-00008

University of Texas at Austin (n.d.). Interpreting results. Retrieved from http://ctl.utexas.edu/programs-and-services/test-scanning-services/interpreting-results/

University of Toronto (n.d.) Post-Master's Nurse Practitioner Diploma. Retrieved April 25, 2009 from http://bloomberg.nursing.utoronto.ca/ prospectives/PosterMaster/ Program_Requirement.htm

van Meter, K. (1990). The collection and interpretation of data from hidden populations. Monograph of National Institute on Drug Abuse, 98. Retrieved from http://archives. drugabuse.gov/pdf/monographs/98.pdf

Waldner, M., & Olson, J. (2007). Taking the patient to the classroom: Applying theoretical frameworks to simulation in nursing education. *International Journal of Nursing Education Scholarship, 4*(1), 1-14. doi:10.2202/1548-923X.1317

West, D., Pomeroy, J., Park, J, Gerstenberger, E., & Sandoval, J. (2000). Critical thinking in graduate medical education: A role for concept mapping assessment? *Journal of the American Medical Association, 284*(9), 1105-1110. doi:10.1001/jama.284.9.1105

Wilgis, M., & McConnell, J. (2008). Concept mapping: An educational strategy to improve graduate nurses' critical thinking skills during a hospital orientation program. T*he Journal of Continuing Education in Nursing, 39*(3), 119-126. doi:10.3928/00220124-20080301-12

Wolf, L. (2005). An ED orientation-fellowship: An experiment in process-oriented training and learning to think like an emergency room nurse. *Journal of Emergency Nursing, 31*(3), 298-301. doi: 10.1016/j.jen.2005.04.018

Wong, F. (2006). Regional orientation program for the department of clinical neurosciences. *Journal for Nurses in Staff Development*, *22*(5), 254- 259. doi:10.1097/00124645-200609000-00009

Woolery, L. (1990). Expert nurses and expert systems: Research and development issues. *Computers in Nursing, 8*(1), 23-28. Retrieved from http://ovidsp.tx.ovid.com.libaccess.lib.mcmaster.ca/sp-3.4.2a/ovidweb.cgi?QS2=434f4e1a73d37e8c7d142c752735f5034bcf1dc368b7aab115d7b1380028dada29298b8d3000ce27b74bde06a410ff5e80246a3917f19e47b33a8465f0da51b

Wray, K., Whitehead, T., Setter, R., & Treas, L. (2006). Use of NCLEX preparation strategies in a hospital orientation program for graduate nurses. *Nursing Administration Quarterly, 30*(2), 162-177. Retrieved from http://ovidsp.tx.ovid.com.libaccess.lib.mcmaster.ca/sp-3.4.2a/ovidweb.cgi?QS2=434f4e1a73d37e8c7cd4188d6221304d1246bd7ba33698af3072

Yoshida, M. (2008). Think-aloud protocols and type of reading task: The issue of reactivity in L2 reading research. In M.Bowles, R. Foote, S. Perpinan, & R. Bhatt (Eds.), *Selected Proceedings of the 2007 Second Language Research Forum* (pp.199-209). Somerville, MA. Retrieved from http://www.lingref.com/cpp/slrf/2007/paper1745.pdf

Zekonis, D., & Gantt, L. (2007). New graduate nurse orientation in the emergency department: Use of a simulation scenario for teaching and learning. *Journal of Emergency Nursing, 33*(3), 283-285. doi:10.1016/j.jen.2007.01.009

Appendix A
Email Recruitment Script

Dear [respondent name],

I am a Master's student at McMaster University School of Nursing with a background in cardiac nursing. I am writing about a research study that I am conducting which explores how cardiac nurses assess patients in clinical situations. The goal of this study is to provide direction to help shape future cardiac nursing professional education programs.

Thank you for agreeing to your colleague's request to contact me regarding your potential participation in my research study. Please consider participating, if any of the following descriptions apply to you:

1. You are a **newly graduated nurse** who:**:**
    - Works in an acute, inpatient adult medical-surgical setting
    - has worked for three months or less in your first nursing position.

OR

2. You are a **registered nurse** who works in one of the following dedicated cardiac settings ("Dedicated" means that cardiology or cardiac surgery is not combined with another service; for example, a combined CCU/ICU is not a dedicated CCU)**:**

i) an inpatient adult cardiac surgery unit,
ii) an adult cardiology unit,
iii) an adult coronary care unit, or
iv) an adult cardiac surgery intensive care unit

AND

- You have two to three years' experience in your current position and you have not worked in any other nursing position.

OR

3. You are an **Acute Care Nurse Practitioner (ACNP)** with**:**
    - at least ten years' experience in adult cardiac nursing, including your experience as a registered nurse prior to becoming an ACNP.

Your participation would consist of the completion of a) a short demographic questionnaire and b) a triad task exercise where you will be asked to compare two very short clinical vignettes against a third "target" vignette. There are a series of ten such triads of vignettes. <u>There are no correct or incorrect answers in this exercise</u>. It should take about 30 minutes of your time. The materials would be sent to you in the mail. In appreciation of your participation, you would receive a $10 Tim Horton's gift card.

Please see the attached consent and information sheet for further details about the study. If you are interested in participating in the study, please respond to this email message and provide your mailing address.

In addition, if you know colleagues who could be eligible for this study, please consider referring them to me via my email address, binkssd@mcmaster.ca. Thank you very much for considering my request, and I hope to hear from you.

Appendix B
Introductory Letter

{Date}
{Address of participant}
Dear {Name of participant},

Thank you for agreeing to participate in this study of differences in the way that nurses at different levels of experience may perceive cardiac patient care situations. Please find enclosed:

a. An information and consent sheet
b. A brief demographic questionnaire
c. Ten triad tasks, with instructions for completion
d. A sheet on which to record your responses
e. A stamped, return envelope

Once you have completed the demographic questionnaire and the triad tasks, please return the signed information and consent sheet, the demographic questionnaire, and the response sheet. Upon receipt of your completed study materials, I will send you a thank you letter, a copy of your signed consent form for your own records, and a Tim Horton's gift card in the amount of ten dollars as a small token of appreciation for your contribution to this study and to the body of nursing knowledge.

Thank you again for your valuable contribution to the study,
Sincerely,

{Signature}
Sally Binks, RN BScN MSc (student) CNCC(C)
School of Nursing,
McMaster University

Appendix C

8/07/09 Protocol # 09-297-S Vers. # 1 Page 1 of 4

McMaster University

Faculty of Health Sciences
School of Nursing
1200 Main Street West
Hamilton, Ontario, Canada
L8N 3Z5
Phone 905.525.9140
Fax 905.524.5228
http://www.fhs.mcmaster.ca/nursing

INFORMATION AND CONSENT SHEET

A Comparison of Nurses' Similarity Judgments at Three Points in the Novice-Expert Continuum

Dear

You are being invited to participate in a research study to explore differences in the way that new and experienced cardiac nurses may assess patients. Research in other domains has shown differences in the perceptions that novices, intermediates, and experts have of problems within that domain. These differences may have important implications for nursing education. A triad task method, described below, will be used to study these differences.

The study is being conducted by Sally Binks, who is a student in the Master of Science (Nursing) program at McMaster University in Hamilton, Ontario working under the supervision of Dr. Ruta Valaitis.

WHO IS ELIGIBLE FOR THE STUDY?

You are eligible to participate in the study if:

1. You are a newly graduated nurse who has worked in any acute adult, inpatient setting for less than three months

   OR

2. You are a registered nurse who:

   a. Works in a dedicated, acute adult inpatient cardiac setting: a cardiology unit, a cardiac surgery unit, a coronary care unit, or a cardiac surgery intensive care unit. (A dedicated cardiac setting is one that is not combined with another service; e.g., a cardiology unit combined with general medicine would not be a dedicated cardiac unit.)

   b. Has worked only in this unit

   c. Has worked in this unit for two to three years.

   OR

3. You are an Acute Care Nurse Practitioner who:

   a. Works in a dedicated, acute, adult inpatient cardiac setting: a cardiology unit, a cardiac surgery unit, a coronary care unit, or a cardiac surgery intensive care unit. (A dedicated cardiac unit is one that is not combined with another service; e.g., a cardiology unit combined with general medicine would not be a dedicated unit.)

School of Nursing

b. Has at least ten years' experience in adult cardiac nursing (including your experience as a registered nurse prior to becoming an ACNP)

It is hoped that 20 nurses from each of these three groups will participate in the study.

WHAT IS EXPECTED OF ME?
Your participation in this study is completely voluntary.
If you agree to be a part of this study, you will be asked to:

1) Complete a 5 minute questionnaire that asks for basic demographic information.
2) Complete a series of ten triad tasks. For each task, you will be presented with three short clinical vignettes. One of the vignettes is a "target" vignette. From the remaining two vignettes, you are asked to select the one that you feel is more like the target vignette and to state briefly the reason for your choice. There are no right or wrong answers in this exercise. Detailed instructions are provided. This activity should take about 30 minutes of your time.

ARE THERE ANY BENEFITS?
Very little is known about differences in the way that nurses at different levels of experience perceive clinical situations. Your participation will help to build knowledge about this. Study results may help to develop effective educational programs for nurses in the future.

ARE THERE ANY RISKS?
It is not expected that there is any risk to you posed by your participation in this study. Participation will in no way affect your employment status or licensure.

WILL I BE PAID TO PARTICIPATE IN THIS STUDY?
You will receive a Tim Horton's Gift Card in the amount of $10 in appreciation of your participation.

CONFIDENTIALITY AND CONSENT
The information that you provide will be shared only with the investigator and her thesis supervision committee. To protect your confidentiality, all responses will be marked with an ID number and no identifying information will be reported or published. Completed hard copies of study materials will be kept in a locked cabinet in an office in the Health Sciences Centre at McMaster University. Data without identifying information will be entered into a software program for analysis on a password-protected computer at McMaster University. Results will be reported as grouped data. If at any time you decide that you do not wish to participate, you can withdraw from the study. There will be no negative consequences to you. Even if you complete the activities, you can later request to have your information excluded from the study. The study data will be destroyed one year after study completion.

HOW CAN I CONTACT SOMEONE IF I HAVE QUESTIONS?
For additional information about the study, you can contact the investigator:
Sally Binks at McMaster University, 905-525-9140 Extension 21422 or via email at binkssd@mcmaster.ca

Or the thesis supervisor:
Dr. Ruta Valaitis, McMaster University, 905-525-9140 Extension 22298 or via email at valaitis@mcmaster.ca

Participants who wish to know the results of the study are invited to contact the investigator.

If you have any questions regarding your rights as a research participant you may contact the Office of the Chair of the Hamilton Health Sciences/ Faculty of Health Sciences Research Ethics Board at 905– 521-2100 ext. 42013.

Hamilton Health Sciences/McMaster
JUL 0 7 2009
Research Ethics Board

CONSENT STATEMENT

I have read and I understand this information and consent sheet. I have had the opportunity to ask questions, and any questions that I have had have been answered to my satisfaction. I agree to participate in this study. I will receive a signed copy of this form for my own records.

____
Name of participant

__________________________ __________
Signature of participant Date

SIGNATURE OF INVESTIGATOR

In my judgement the participant is knowingly and voluntarily giving consent to participate in this study

__________________________ __________
Name and signature of investigator Date

Appendix D

DEMOGRAPHIC QUESTIONNAIRE

1. **I currently work on:**
   - ☐ A dedicated inpatient adult cardiology unit
   - ☐ A dedicated inpatient adult cardiac surgery unit
   - ☐ A dedicated adult coronary care unit
   - ☐ A dedicated adult cardiac surgery intensive care unit
   - ☐ An acute inpatient, adult medical-surgical unit.

   Please note: For the purposes of this study, a "dedicated" unit is defined as one that is not combined with another service; e.g., a combined ICU/CCU is not a dedicated unit.

2. **I am a(n) (select all that apply):**
   - ☐ Registered nurse
   - ☐ Acute Care Nurse Practitioner
   - ☐ A graduate nurse working under a temporary license

3. **My highest level of education in nursing is:**
   - ☐ Diploma
   - ☐ Baccalaureate degree
   - ☐ Master's degree
   - ☐ PhD degree

4. **I have worked in my current role (staff nurse or nurse practitioner) for _____ months or _____ years.**

5. **I have worked in my current position for ____ months or ______ years.**

Appendix E

Final Study Instrument

MATCHING TRIAD TASK

For each of the following questions, you will see a set of three short descriptions; one description (A) is positioned above the two other two descriptions (B and C). From B and C, please select the description that you believe is most similar to description A and write a statement to indicate your reason for making that choice on the response sheet. Please indicate only one choice-the choice you consider most important. ***Describe in as much detail as you can the rationale for your choice*** Remember, there are no right or wrong answers.

1.

MK is a 67-year-old man with congestive heart failure who receives furosemide (Lasix).

A

EM is a 69-year-old man with congestive heart failure who receives ramipril (Altace).

B

JH is a 55 year-old woman with angina who wears a nitroglycerin patch (Nitro-Dur).

C

2.

KF is a 33 year-old woman with a diagnosis of septic shock secondary to endocarditis. Her blood pressure is 80/46

A

JL is a 30-year old man with anaphylactic shock secondary to peanut allergy. His blood pressure is 77/40 mm Hg.

B

OP is a 40-year old man who is in cardiogenic shock secondary to anterior wall MI. His blood pressure is 84/56 mm Hg.

C

3.

RV is a 70-year-old man who has had an inferior wall MI. He has symptomatic bradycardia and has had a temporary pacemaker inserted.

A

KL is an 80-year-old woman who has gone into rapid atrial fibrillation and receives amiodarone (Cordarone) 300 mg IV.

B

WB is a 68 year-old man who developed mitral regurgitation after an inferior wall MI.

C

4.

AI is a 70 year-old woman who has had mitral valve surgery. She has had a permanent pacemaker inserted for persistent post-operative third-degree AV block.
A

ED is a 77 year-old man who has had aortic valve replacement for aortic stenosis.
B

ST is a 66 year-old man who takes digoxin (Lanoxin) for chronic atrial fibrillation.
C

5.

AI is a 70-year-old woman who has had a permanent pacemaker implanted for persistent post-operative bradycardia after mitral valve replacement.
A

ED is a 77-year old man who has had aortic valve replacement for aortic stenosis.
B

TP is a 19-year-old man who has had ablation therapy for a re-entry tachycardia.
C

6.

HB has pulmonic stenosis.

A

KE has pulmonary hypertension.

B

RI has aortic stenosis.

C

7.

JK has CHF and takes captopril (Capoten).

A

PL has CHF and takes digoxin (Lanoxin).

B

LN has hypertension and takes hydralazine (Apresoline).

C

8.

DM is a 83 year-old man who receives blood products after a mitral valve replacement.

A

JH is a 77 year-old woman who receives IV nitroglycerin after an aortic valve replacement.

B

EM is a 17 year-old girl who receives a normal saline infusion for diabetic ketoacidosis.

C

9.

PK has a temporary pacemaker that is oversensing.

A

LP has third-degree block with a ventricular escape rhythm.

B

RC has a temporary pacemaker that is undersensing.

C

10.

DH has septic shock.	A

OB has cardiogenic shock.	B

WS has severe anemia.	C

Appendix F

RESPONSE SHEET

1. ___ is more like A because ____
____.
2. ___ is more like A because ____
____.
3. ___ is more like A because ____
____.
4. ___ is more like A because ____
____.
5. ___ is more like A because ____
____.
6. ___ is more like A because ____
____.
7. ___ is more like A because ____
____.
8. ___ is more like A because ____
____.
9. ___ is more like A because ____
____.
10. ___ is more like A because ____
____.

Thank you for your participation!

Appendix G

Thank You Letter

{Date}

{Address of participant}

Dear {Name of participant}

Thank you for taking the time to complete the study materials for my Master of Science (Nursing) research project. I hope you found the triad task exercise interesting. This work has not only advanced my own learning tremendously, but I hope that it will contribute to the body of knowledge that will inform specialty nursing education in the future. As a small token of my appreciation for your participation in this study, please find enclosed a Tim Horton's gift card in the amount of $10.

Sincerely,
{Signature}
Sally Binks, RN BScN MSc (student) CNCC(C)

Appendix H

Clarity Evaluation

Readability and Clarity Evaluation
Instructions:

- First, complete the attached demographic questionnaire and the series of ten Matching Triad Tasks; instructions are included at the top of the first page
- Please answer the questions below.

1. Time to complete the demographic questionnaire ____________minutes.
2. Time to complete Matching Triad Tasks __________________minutes.
3. Are the demographic questionnaire items clear and understandable? Please comment below.
4. Are the instructions for completing the Matching Triad Tasks clear and understandable? Please comment below.
5. Do you understand all of the words used in the demographic questionnaire? If no, please indicate the words that were not understood.
6. Do you understand all of the words used in the instructions for completing the Matching Triads Tasks and in the Matching Triads Tasks? If no, please indicate the words that were not understood.
7. Does the questionnaire formatting enhance the clarity and ease of completion of the questionnaire? Please comment below.
8. Does the formatting of the Matching Triad Tasks enhance the clarity and ease of completion? Please comment below.
9. General impression:
10. Other comments

Appendix I

Revised Forced Choice Triad Task

MATCHING TRIAD TASK

For each of the following questions, you will see a set of three short descriptions; one description (A) is positioned above the two other two descriptions (B and C). From B and C, please select the description that you believe is most similar to description A and write a statement to indicate your reason for making that choice on the response sheet. Please indicate only one choice-the choice you consider most important. ***Describe in as much detail as you can the rationale for your choice*** Remember, there are no right or wrong answers.

1.

MK is a 67-year-old man with congestive heart failure who receives furosemide (Lasix).
A

EM is a 69-year-old man with congestive heart failure who receives ramipril (Altace).
B

JH is a 55 year-old woman with angina who wears a nitroglycerin patch (Nitro-Dur).
C

2.

KF is a 33 year-old woman with a diagnosis of septic shock secondary to endocarditis. Her blood pressure is 80/46

A

JL is a 30-year old man with anaphylactic shock secondary to peanut allergy. His blood pressure is 77/40 mm Hg.

B

OP is a 40-year old man who is in cardiogenic shock secondary to anterior wall MI. His blood pressure is 84/56 mm Hg.

C

3.

MH is a 55-year-old woman who had a large anterior wall MI three days ago. Her heart rate is 110/min. Her BP is 92/58 mm Hg

A

JT is a 60-year-old man who has haemorrhaged after orthopaedic surgery. His BP at 1200 was 105/50 and his HR was 102/min; his BP at 1210 is 100/60 and his HR is 110/min

B

TV is a 50-year-old man who has had bypass surgery. He rates his incisional pain at 9 out of 10. His heart rate is 110/min and his BP is 130/72 mm Hg

C

4.

RV is a 70-year-old man who has had an inferior wall MI. He has symptomatic bradycardia and has had a temporary pacemaker inserted.
A

KL is an 80-year-old woman who has gone into rapid atrial fibrillation and receives amiodarone (Cordarone) 300 mg IV.
B

WB is a 68 year-old man who developed mitral regurgitation after an inferior wall MI.
C

5.

EL is a 80-year-old woman who has congestive heart failure. She suddenly develops rapid atrial fibrillation. Her extremities are cold and pale.
A

MP is a 78-year old man who suffers from peripheral arterial disease. His feet are cold and pale.
B

JK is a 45 year-old man who has had CABG X4 and diuresed two litres in an hour. His extremities are cold and pale.
C

6.

IK is an 80 year-old man with pulmonary hypertension.

A

WL is a 48 year-old man with hypertension.

B

CW is a 50 year-old woman with aortic stenosis.

C

7.

AI is a 70 year-old woman who has had mitral valve surgery. She has had a permanent pacemaker inserted for persistent post-operative third-degree AV block.

A

ED is a 77 year-old man who has had aortic valve replacement for aortic stenosis.

B

ST is a 66 year-old man who takes digoxin (Lanoxin) for chronic atrial fibrillation.

C

8.

AI is a 70-year-old woman who has had a permanent pacemaker implanted for persistent post-operative bradycardia after mitral valve replacement.
A

ED is a 77-year old man who has had aortic valve replacement for aortic stenosis.
B

TP is a 19-year-old man who has had ablation therapy for a re-entry tachycardia.
C

9.

CV is a 48-year-old woman who has developed cardiac tamponade after CABG.
A

JB is a 62-year-old man who has constrictive pericarditis.
B

MW is a 50-year-old woman who develops a clot in her mechanical mitral value.
C

10.

RF is a 68-year-old woman with neurogenic shock.
A

KV is an 80-year-old man with cardiogenic shock.
B

TG is a 45-year old man with septic shock.
C

11.

HB has pulmonic stenosis.
A

KE has pulmonary hypertension.
B

RI has aortic stenosis.
C

12.

JK has CHF and takes captopril (Capoten).
A

PL has CHF and takes digoxin (Lanoxin).
B

LN has hypertension and takes hydralazine (Apresoline).
C

13.

DM is a 83 year-old man who receives blood products after a mitral valve replacement.
A

JH is a 77 year-old woman who receives IV nitroglycerin after an aortic valve replacement.
B

EM is a 17 year-old girl who receives a normal saline infusion for diabetic ketoacidosis.
C

14.

UB has acute mitral regurgitation.
A

GH has an acute exacerbation of congestive heart failure.
B

CR has a small congenital ventricular septal defect.
C

15.

PK has a temporary pacemaker that is oversensing.
A

LP has third-degree block with a ventricular escape rhythm.
B

RC has a temporary pacemaker that is undersensing.
C

16.

DH has septic shock.
A

OB has cardiogenic shock.
B

WS has severe anemia.
C

Appendix J

Content Validity Assessment

This research study will investigate possible differences in the way that novices and experts perceive clinical situations: The hypothesis is that novice nurses will tend to focus on the "surface" aspects of a given situation, while those in the expert cohort will tend to perceive the "deep" principles of the same situation. The hypothesis will be tested by means of a methodology called "forced choice triad task". Each task consists of three brief items: a "target" item (item A) and two referent items (Items B and C). Participants will be asked to select the referent item (either B or C) that they believe is more similar to the target item (A), and to indicate the reason for their choice. In each triad, one referent item was designed to have surface characteristics, such as patient age and sex or diagnosis, that are similar to the target item, while the other referent item was designed to be similar to the target on the basis of deep features. The deep similarities between referent and target items represent three categories that reflect principles of hemodynamics: 1) manipulation of the determinants of cardiac output (preload, contractility, afterload and heart rate) to achieve a therapeutic goal; 2) the operation of cardiovascular compensatory mechanisms (such as reflex increase in heart rate to compensate for a decline in stroke volume) in the maintenance of adequate cardiac output and perfusion; and 3) the abnormality in the determinants of cardiac output that underlies a particular pathophysiology (such as excessively low afterload as the basic pathophysiologic mechanism in certain shock states).

In order for research findings to be valid, a research instrument must accurately and completely measure or assess what it is intended to measure. As a content expert in cardiac nursing, your opinion of the accuracy and completeness of this instrument is important. There are two parts to this assessment: Part I requests your assessment of the degree to which each item reflects the basis of surface or deep matching on which it was designed; Part II requests your opinion of how completely the instrument reflects the principles of hemodynamics that nurses need to understand in order to provide safe care. As you complete Part I, please refer to the chart below, which describes the basis on which the deep and surface matches were designed for each referent item; for example, for Triad Task #1, referent item B was designed to be similar to the target item (A) on the basis of surface features (patient age, sex, diagnosis, and the fact that a drug is mentioned), and referent item C was designed to be similar to the target item on the basis of a surface feature (the mention of drug therapy) and a deep feature (the use of drug therapy to therapeutically reduce preload).

Surface and Deep Similarities of Referent Items to Target Item

Task	Referent Item B Surface Similarities to Target	Referent B Deep Similarities to Target	Basis of Deep Similarity	Referent C Surface Similarities to Target	Referent C Deep Similarities to Target	Basis of Deep Similarity
1	Age, sex, diagnosis, mention of drug therapy.			Mention of drug therapy	Drug therapy to reduce preload	Therapeutic goal
2	Hypotension. Shock state. Age	Excessively low afterload	Basic physiologic problem	Hypotension Cardiac problem Shock state.		
3	Abnormal vital signs Age	Operation of compensatory mechanisms: tachycardia and increased diastolic BP.	Operation of compensatory mechanisms	Age Cardiac problem Tachycardia		
4	Arrhythmia	Manipulation of heart rhythm/rate to improve cardiac output.	Therapeutic goal.	Age, sex, cardiac problem.		
5	Coldness, pallor of extremities. Age.			Coldness, pallor of extremities.	Selective vasocon-striction to maintain vital organ perfusion.	Operation of compensa-tory mechanisms
6	Hypertension.				High RV afterload	Basic physiologic problem.
7	Valve surgery.				Control of rhythm to maintain CO.	Therapeutic goal.

Task	Referent Item B Surface Similarities to Target	Referent B Deep Similarities to Target	Basis of Deep Similarity	Referent C Surface Similarities to Target	Referent C Deep Similarities to Target	Basis of Deep Similarity
8	Age, Valvular surgery.				Manipulation of heart rate or rhythm to improve cardiac output.	Therapeutic goal.
9		Increased resistance to diastolic filling	Basic physiologic problem	Age, sex, cardiac surgery.		
10	Shock state			Shock state	Excessively low afterload	Basic physiologic problem
11	Pulmonary problem	Excessively high RV afterload	Basic physiologic problem	Valvular stenosis		
12	Drug therapy			Drug therapy	Therapy to reduce afterload.	Therapeutic goal.
13	Age Valvular surgery				Volume expansion to maintain CO	Therapeutic goal.
14	Valvular stenosis				Conditions that may cause right-sided failure	Basic physiologic problem.
15		Heart rate excessively slow.	Basic physiologic problem.	Pacemaker; sensing problems.		
16	Shock states				Conditions resulting in abnormally high CO	Basic physiologic problem.

PART I

For each referent item, please indicate on the four-point scale how poorly or how well you believe each referent item reflects the basis on which it was designed. Please feel free to add comments to clarify your rating, or to comment on the appropriateness or clarity of the items.

1. **Triad Task #1.**

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

2. **Triad Task #2.**

a) *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

3. **Triad Task #3**

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

4. **Triad Task #4**

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

5. Triad Task #5

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

6. Triad Task # 6

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

7. Triad Task # 7

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

8. Triad Task #8

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

9. Triad Task #9

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

10. Triad Task # 10

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

11. Triad Task # 11

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

12. Triad Task # 12

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

13. Triad Task # 13

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

14. Triad Task # 14

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

15. Triad Task # 15

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

16. Triad Task # 16

a). *Item B* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

b). *Item C* is similar to *Item A* on the basis of the features indicated in the chart.

1 Strongly disagree	2 Disagree	3 Agree	4 Strongly agree.

PART II

For the instrument as a whole, please comment on the degree to which you believe this instrument reflects the principles of hemodynamics that nurses in clinical practice need to understand in order to provide safe care. Are any important concepts omitted? Please be as clear and precise as you can.

____
____
____
____

THANK YOU FOR YOUR VALUABLE INPUT