Thinking Like a 21st Century Nurse: Theory, Instruments, and Methodologies for Measuring Clinical Thinking

Lily Fountain
University of Maryland

Please address all correspondence to:
Lily Fountain
University of Maryland
School of Nursing
655 W. Lombard St. Room 590D
Baltimore, MD 21201
410.706.1763 office
410.706.0253 fax
fountain@son.umaryland.edu

Paper presented at the
Annual Meeting of the American Educational Research Association
New Orleans, April 8-12, 2011

Thinking Like a 21st Century Nurse: Theory, Instruments, and Methodologies for Measuring Clinical Thinking

Abstract: This cross-sectional descriptive study of the Model of Domain Learning, which describes learners' progress from acclimation through competence to proficiency through the interplay of knowledge, interest and strategic processing/critical thinking (CT), examined its extension to maternity nursing. Based on the identified need for valid, reliable quantitative instruments measuring cognitive and affective aspects, three instruments were developed: a 20-item, polytomously-scored multiple choice questionnaire, a five item Interest Survey, and a written CT case scenario analysis. The sample was 87 baccalaureate student nurses in the third and final semesters. The instruments demonstrated mixed support for the Knowledge, Interest, and CT scales. Three principal component factors mapped well onto current definitions of CT. Further refinement of instruments and a broader sample were recommended.

The complexity of the current health care system has placed increasing demands on health professional education. Patients are sicker, older, and more culturally diverse, and the structure of the health care system is constantly fluctuating due to changes in insurance, regulations, and technology. An understanding of the trends making demands on professional education will improve the application of theories, instruments, and methodological solutions.

Regarding trends in patient care, patient classification systems have indentified increases in such measures as the average case mix index (Jennings, 2008) that indicate a more complex caseload for nursing care. Technologies used in the care of patients such as pumps, robots, medication delivery systems, computers and documentation systems, are changing every day, and increasing consumerism in patients has added a new dimension to patient teaching (Cohen, Grote, Pietraczek, & Laflamme, 2010). Another trend that is increasing the complexity of care is the aging of the U.S. population, with an increasingly diverse racial and ethnic composition (Jacobsen, 2011). The demographics of nursing students themselves are changing as the profession becomes more racially, ethnically, internationally, and socioeconomically diverse, with increased gender and age distribution (AACN, 2008).

The new Health Care Reform laws and regulations will require nurses to care for patients more safely, accurately, and in a manner that utilizes evidence-based practice. The new programs will utilize more community-based settings where access to experienced mentors may be decreased (AACN, Apr. 2010). The quality assurance demands via audit increase every day, as the cost and efficiency of care delivery are scrutinized more closely (RWJF, Dec. 2008).

Regarding trends in nursing education, there have been widespread professional calls for improvements in the education of nurses. The Institute of Medicine (IOM) is an independent non-profit that is an arm of the National Academy of Sciences that serves as a national advisor on health. Its recent report, *The Future of Nursing(2008),* calls for increases in decision-making skills of nurses in educational programs. The Carnegie Foundation for the Advancement of Teaching recently released *Educating Nurses* in the Preparation for the Professions series, which recommends that nurse educators emphasize clinical reasoning that incorporates the many factors that must be considered in providing nursing care (Benner, Sutphen, Leonard, and Day,

2010). The American Association of Colleges of Nursing released the *Baccalaureate Essentials* in 2008 that provided a framework for baccalaureate nursing education that emphasized "clinical reasoning/critical thinking" as well as other concepts (AACN, 2008).

With strong societal and professional pressures and with nearly a half million baccalaureate nursing students in the US (AACN, 2008) and 3.1 million practicing Registered Nurses, there is a large responsibility for nursing faculty to safeguard and improve the quality of thinking among nursing students and practicing nurses. These trends have increased the need for teaching strategies for improving critical thinking in the incoming nursing workforce and for measures that can evaluate critical thinking. Evaluating critical thinking requires theories that are robust enough to explain individual and cohort changes, instruments that are precise enough to capture components of professional practice yet generalizable enough to be used in different clinical settings, and methodologies that capture nuances in performance data.

Definitions

A review of the literature on critical thinking in nursing education reveals the following themes in the research: focus on the definition of critical thinking and related concepts in order to capture all aspects of nursing practice, and the use of standardized and researcher-developed instruments,.

The initial impetus for increased study of thinking in nursing came from nursing program accreditation requirements for nursing education programs to demonstrate critical thinking (CT) in curricular outcomes in 1991 (Simpson & Courtney, 2002). Much professional discourse has been spent on defining critical thinking. In 1990, an APA Consensus Panel led by Facione defined CT as "purposeful, self-regulatory judgment, which results in interpretation, analysis, evaluation, and inference, as well as explanation of the …considerations on which that judgment is based" (Facione, 1990, p.2). In the mid 1990's Scheffer and Rubenfeld conducted a three year Delphi study to gain consensus from a diverse group of expert nurses using a process similar to the APA process. They identified 7 cognitive strategies and 10 dispositions or habits of mind that have been used by many nursing researchers: the skills of analyzing, applying standards, discriminating, information seeking, logical reasoning, predicting, and transforming knowledge, as well as the dispositions or "habits of mind" of confidence, contextual perspective, creativity, flexibility, inquisitiveness, intellectual integrity, intuition, open-mindedness, perseverance, and reflection (Scheffer & Rubenfeld, 2000a). There were a great number of similarities in the characteristics identified by both groups. Of note, creativity, intuition, and transforming knowledge were identified for nursing but not identified by the APA group.

At least 11 other definitions of CT are published in the nursing literature (Tanner, 1983; Itano, 1989; Facione, 1990; Jones and Brown, 1991; Kataoka-Yahiro & Saylor, 1994; Oermann, 1997; Walsh & Seldomridge, 2006; Walters, 1986; Alfaro-LeFevre,1999; Daly, 1998;Edwards, 2006), although there is little evidence of attempts to build upon previous definitions in a consistent fashion. Both the cognitive and dispositions/affective aspects of CT have been explored in the literature.

Related concepts

Several terms are often used interchangeably with CT: problem solving, decision making, and clinical judgment. Some distinctions between the terms can be made, but often the most important difference is the different paradigms or research literatures that the terms are used in. Overlap still occurs. Problem solving is often cited as a synonym for critical thinking. However, problem-solving is focused on a specific outcome or solution, whereas CT looks at the larger picture, and sometimes more ill-structured problems (Simpson & Courtney, 2002). Problem solving is closely related to Information Processing approach to cognitive development, with an emphasis on cue acquisition and interpretation and hypothesis generation and evaluation (Roberts, 2000).

Another term frequently used synonymously with CT is clinical decision-making. Decision-making focuses on the specific behavior that nurses must perform: whether to turn on the oxygen, whether to administer a drug. Clinical judgment or reasoning requires experience of many patient cases to develop over time. Much of the research in nursing and medical practice relating to these terms uses the novice/expert paradigm, and much of the research is based on medical education research. Although medicine and nursing both deal with health and illness and patients, they are completely different professions and require different constructs, methodologies, and teaching strategies to some extent. For instance, correct medical diagnosis is paramount in medicine, whereas in nursing, the focus is on the patient/client's response to illness. An important profession-specific finding is that the process by which nurses deliver care to patients, the nursing process, is not considered equivalent to CT by most authors (Brunt, 2005; Kataoka-Yahirio & Saylor, 1994). The stages of the nursing process, assessment, planning, nursing diagnosis, intervention, and evaluation, do not include the cognitive strategies such as inferences and finding arguments that are part of CT, nor does the nursing process address the habits of mind needed in CT such as inquisitiveness and reflection. Some scholars view CT, problem solving, decision-making and clinical judgment as multiple types of thinking strategies that are all needed by nurses in different situations for high quality practice (Benner, Hughes, & Sutphen, 2008). In addition, the relationship-based and patient-centered aspects of care are not captured by some definitions (Tanner, 1997). Although most nursing studies focused on the construct of critical thinking, the bodies of research on clinical reasoning and problem solving offer techniques and instruments that operationalize critical thinking as utilized in nursing education literature

Strategic processing

The related concept of strategic processing has also been studied in education literature. Strategic processing refers to the use of strategies to acquire, organize, and transform information (Samuelstuen & Braten, 2007). In a study of the relationship of critical thinking, motivation, and classroom experiences, deep processing strategies such as elaboration and metacognition were found to be correlates of critical thinking (Garcia & Pintrich, 1992, p. 5). Strategic processing has been studied in nursing (Braten and Olaussen (2007). In a longitudinal study of motivation in nursing students , the authors found that the more positively motivated students were found to report more use of not only deeper but also surface processing strategies such as memorization. However, the use of deep processing strategies decreased from the first

year to the second, although the use of superficial strategies stayed the same. The authors hypothesized that nursing schools may give undue rewards for rote memorization in tests and other assignments. A limitation on generalizability to the US was that the study took place in Norway, and it is not known how similar the Norwegian nursing curriculum is to that of the US. Educational researchers have found that memorization results in short-term learning (Pintrich et al, 1991), whereas deep processing strategies seem to promote longer-term retention (Weinstein et al., 2000). Alexander (2004) has found that superficial processing decreases over the course of professional development.

Standardized Instruments

Due to a lack of consensus on the definition of critical thinking, and due to accreditation requirements to demonstrate assessment of critical thinking, many nursing schools use standardized instruments to measure CT (Brunt, 2005; Facione & Facione, 1994). Standardized tests found during this review in the nursing education literature included the California Critical Thinking Skills Test (CCTST) and California Critical Thinking Skills Disposition Inventory (CCTSDI), the Watson Glaser Critical Thinking Skills Appraisal (WGCTSA), the Enis-Weir Critical Thinking Essay Test, and the ERI Critical Thinking Process Test (CTPT). The Cornell Critical Thinking Test was mentioned but no other data was located (Oermann & Gaberson, 1998).

The Watson-Glaser Critical Thinking Skills Appraisal, WGCTA, revised in the 1980's (Facione & Facione, 1994) has been widely used on college students, as well as by nursing schools and has 80 items, with two versions. It is a multiple choice test with 5 subtests with 16 items each: Inference, recognition of assumptions, deduction, interpretation, and evaluation of arguments. It is not specific to any domain, and does not capture the affective dimensions of CT. Studies using this instrument to assess change in CT as measured by the WGCTA over the course of the nursing program typically found no change or a decrease in CT (e.g. Daly, 2001; Walsh & Seldomridge 2006b). Complaints from researchers using the instrument included: pre-licensure is too soon to measure CT,; CT needs to be taught more explicitly in nursing programs; nursing-specific instruments need to be developed; the teaching of the CT skills measured by the instrument such as logical reasoning in the nursing program; CT skills be divided into skills that novices could expect to increase and ones that more experienced nurses would use more often, such as creativity.

Educational Resources International, Inc. developed a CT test called the Critical Thinking Process Test (CTPT). It is a composite of 5 scales, Prioritizing, Reasoning, Goal Setting, Application, and Evaluating. Hoffman (2006) found a statistically significant increase in CT as measured by the CTPT from the beginning to the end of the nursing program among three cohorts of students, with a total N of 437. The study is notable for the large N and control for many variables in a multivariate analysis. In ERI's own studies they found CT as measured by CTPT increased over the course of the nursing programs. This instrument is not widely used and is expensive to administer.

The most widely used instruments are the California Critical Thinking Skills Test (CCTST) and the related California Critical Thinking Dispositions Inventory (CCTSDI). The CCTST is a 34-item instrument designed to measure CT in college-age students, based on the APA Delphi study. The iterative Delphi process used as a basis for the study was described above. The CCTST assesses areas similar to the WGCTSA, including the cognitive skills of analysis, evaluation, inference, inductive, and deductive reasoning. The California Critical Thinking Skills Disposition Inventory (CCTSDI) has 75 Likert type items and has 8 independent subscale scores: inquisitiveness, open-mindedness, systematicity, analyticity, truth-seeking, CT self-confidence, and maturity (Facione, 1990). It examines changeable "habits of mind" that promote CT. A sample item is "We can never really learn the truth about most things", or "The best argument for an idea is how you feel about it at the moment", or "Advice is worth exactly what you pay for it", with the Likert scale ranging from "strongly agree" to "strongly disagree" (Tishman & Andrade, N.D). Ten studies and one meta-study were located that used this instrument, of which 5 examined entry/exit changes in CT.

McMullen & McMullen (2008) used the CCTST in a longitudinal study of the development of critical thinking in nursing . They found that the student's percentile, 25^{th}, 50^{th}, or 75^{th}, affected the trajectory of growth over the course of the nursing education program, with higher percentile students making slower gains or even decreases compared to low percentile students who increased critical thinking skills. This is the only nursing study found that used a longitudinal design as opposed to pre/post. The authors concluded that standardized tests should not be used for testing for CT, and that CT should be taught explicitly in the curriculum.

In spite of the strong content validity and wide use, results have also been inconsistent with these tests (e.g. Beckie, Lowry, and Barnett 2001). There are two possible explanations for the lack of consistent increase in CT as measured by CCTST/CCTSDI: 1) Nursing education is not promoting critical thinking; 2) the instruments are not valid for this domain.

However, some authors have noted the possibility that nursing curricula are not promoting critical thinking to the extent possible (e.g., Braten & Olaussen, 2007). Walsh and Seldomridge (2006a) examined the types of thinking being reinforced in nursing curricula. They were concerned that the lecture format for teaching, limited class time, multiple choice examinations, publisher-made or pre-packaged power point presentations and administrative pressure to use them, and student expectations for "sage on the stage" entertainment, are all factors that have contributed to superficial thinking in nursing classes.

In efforts to find a theory that can explain clinical thinking in nursing, Tanner (2006) offers a model of clinical judgment (CJ) that is a recursive process of noticing, which includes contextual and patient cues as well as assessment and textbook knowledge, then the nurse pursues one of the analytic processes, and chooses an action, and then reflects on action, or evaluates. A rubric for evaluating clinical thinking according to this model was developed for a nursing simulation (Lasater, 2007). The Oregon Health and Science University School of Nursing faculty team have empirically validated this model and rubric using simulations and clinical evaluation. This is one of the few instances of a program of research relating to the measurement and development of clinical thinking in nursing.

In addition to these cited empirical studies on educational strategies for CT in nursing, a meta-analysis of teaching strategies used in all domains to promote critical thinking by Abrami et al. (2008) found an average effect size of 0.34 (k=161,n=20,698, SD=0.6). Instruction improved CT and dispositions to critical thinking. The greatest effect size among types of interventions was seen with teacher-made CT interventions. The greatest effect size was seen if a "mixed" approach, with subject-specific CT instruction and a separate thread or course aimed at teaching general principles of CT was provided (Effect Size ES=0.94). The second greatest effect was seen with "infusion" instruction, where there was deep subject matter instruction on CT, as well as general principles of CT skills were provided (ES=0.54). Other statistically significant approaches were "immersion", where subject matter-specific CT instruction was provided but CT principles were not made explicit (ES=0.09), and "general instruction", where CT skills and dispositions were learning objectives without subject matter content(ES=0.38) (typology from Ennis, 1989). The effect size did not vary much by type of research design (experimental (ES=0.34), quasi-experimental (ES=0.36), and pre-experimental (ES=0.31). Also important was pedagogical grounding of the faculty in CT; if the instructor had a course in CT, effect size was 1.00; if the instructor had extensive observations, effect size was 0.58; and if the instructor had developed a detailed curriculum description, the effect size was 0.31; if CT was listed as a course objective, effect size was 0.13. Only one nursing study met the criteria for inclusion in this review.

Teacher-Made Instruments

Nursing faculty researchers have designed instruments or surveys to analyze CT when evaluating teaching strategies. The same definitional diversity is seen. No empirical research was present for most of the instruments used to evaluate the CT changes from teaching strategies. These teacher-made instruments have been used to evaluate teaching strategies such as critical incident discussions, joint rounds, paradigm cases, and seminars (Brunt, 2005). Simpson and Courtney (2002) list role-playing, debate, jigsaws, writing assignments, and simulations as teaching strategies purported to increase CT.

Concept maps have been used to measure CT. Although there are studies indicating success in increasing CT through concept maps, (e.g. Abel and Freeze, 2006), instructional challenges include inter-rater reliability, and time required for orientation, administration, and grading. Advantages include that it is a strategy that captures student understanding of relationships, can be used to follow student development, and reliable grading rubrics have been designed (Hsu, 2004).

Gaps Identified by the Literature

Problems have been identified with the definition, measurement, research methods, and educational implementation of CT. Traditional methods of nursing education have not been consistently effective in increasing CT, and some studies have shown a decrease in critical thinking over the course of schooling. It is difficult to unpack if inconsistencies are due to instructional differences or the difficulties in measuring CT.

Definitional diversity and lack of a strong theoretical base for most instruments was a problem, as models of a construct, not just definitions, are needed to correctly operationalize it.

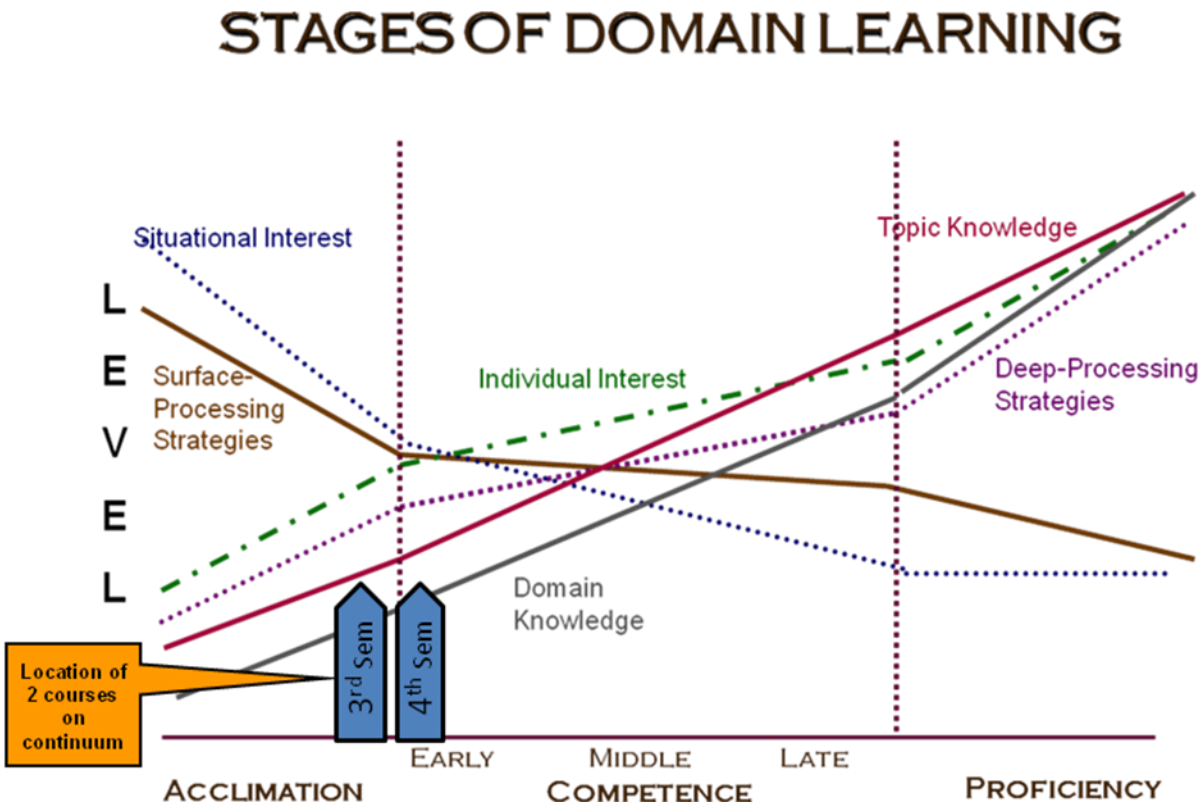
Many researchers called for domain-specific instruments to measure nursing CT. With the exception of Tanner's work (2006) and the standardized CT tests with inconsistent results, few instruments were used for more than one study, and statistical validity and reliability were seldom reported. Many promising instruments remain buried in unpublished dissertations. Another often missing aspect was the measurement of non-cognitive parts of nursing care, such as motivation.

Framework

To address these gaps, this study uses the Model of Domain Learning (MDL) (Figure 1) as a framework for studying the development of nurses from acclimating novice to proficient nurse. The MDL was developed in the context of educational psychology by Alexander (1997) and investigated by Alexander and colleagues (e.g. Alexander 2004). This developmental expertise model has been researched across many domains, and examines the changes in Knowledge, Interest, and Strategic Processing as individuals move from acclimation to proficiency in an academic domain (see Figure 1). This model has several features that make it promising for this nursing research: 1) Strategic Processing captures the surface and deep aspects of critical thinking strategies identified in previous research; 2) the model has an affective component through the Interests construct, and 3) Nursing has seldom examined different types of Knowledge, Interest and Strategic processing. This model explores the dual aspects of domain knowledge and topic knowledge. Both fleeting Situational Interest such as that engendered by an exciting speaker as well as enduring Individual Interest demonstrated by most nurses as they specialize in an area of patient care are characterized. The changing nature of the types of strategies used by learners over their professional course are described by surface strategies such as patient problem description in nursing, to deep processing strategies such as justifying hypotheses (Kamin, O'Sullivan, Younger, & Deterding, 2001).

FIGURE 1.

Model of Domain Learning



Alexander, P. A. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational, and strategic forces. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 213-250). Greenwich, CT: JAI Press.

The two objectives of this pilot study were 1) to determine if a theoretical model and instruments used to explain changes in knowledge, interest, and strategic processing in reading and other academic domains could be extended to a clinical domain such as maternity nursing, and 2) to determine if critical thinking could be objectively measured in a written case scenario format in the domain of maternity nursing.

Method

Sample

For this pilot study to answer these questions, a convenience sample of 87 pre-licensure nursing students from a large Mid-Atlantic University were recruited between 2008 and 2010. This students in this sample were in a "2+2" or upper division entry level nursing program where nursing science prerequisites are completed prior to the last four prelicensure semesters in the nursing program. These four semesters of nursing courses include didactic and clinical components. In Semester One of the program (the Junior year) students complete Fundamentals

of Nursing; during Semester Two students complete the clinical course in Adult Medical Surgical Nursing; and in the Semester Three, they complete the Pediatric, Psychiatric, and Obstetric Nursing clinical courses. The Maternity nursing course includes a seven week 90 hour clinical in basic maternity nursing. In the final Semester Four, students are enrolled in nursing courses in Community Nursing and Senior Practicum-Integration in a specialized area of Nursing: Medical Surgical, Critical Care, Pediatric, Psychiatric, or Obstetric. During this final semester the students apply and integrate the knowledge, skills, and strategies learned in previous semesters to one specialized area of nursing. During this fourth semester students complete 180 clinical hours of practicum in this specialty area of their choice, and 90 hours of Community nursing clinical. The demographics of the sample are shown in Table 1. Of the 87 students, 50 (57%) were at the end of the third semester, and 37 (43%) were at the end of the fourth semester. The sample was 90% female and 10% male, 37% African-American or African, 48 % Caucasian, and 14 % Asian, and 5% reported Hispanic ethnicity. The mean age was 27.6 years, with a SD of 6.0 and a range of 21 to 48 years. There were no statistically significant differences for the demographic variables between the 3rd semester students and the fourth semester students in Obstetrical and Other Specialties.

TABLE 1. *Sample Demographics*

	N=87
Gender	Female 78 (90%) Male 9 (10%)
Age	Mean=27.6 years SD=6.0 N=85 Range=21-48 Missing=2 (2%)
Race	Black/African American 32 (37%) White/European-American 42 (48%) Asian 12 (14%) Missing 1 (1%)
Hispanic	Yes 4 (5%) No 80 (92%) Missing 3 (3%)
Course Level of Student	3rd Semester 50 (57%) 4th Semester Practicum 37 (43%) Practicum Specialty Medical Surgical Nursing 10 (27%) Critical Care Nursing 5 (14%) Obstetrical 7 (19%) Pediatric 12 (32%) Psychiatric/Community 3 (8%)

Recruitment and Procedure

Participants were recruited during Maternity Nursing and Senior Practicum classes. For five semesters during the duration of the study, during class time, the Principal Investigator introduced and described the study and left the room. The Research Assistant reviewed the Research Consent Forms, allowing students time to read and sign the consent form, and collected them. Because the Principal Investigator was a faculty in the two courses in which participants were enrolled, the research sessions were conducted by research assistants who were not faculty to the participants. Research sessions were at a scheduled time determined by final examination time and student schedules. Since nursing students are usually in class when they are on campus this was difficult to arrange. Over the four semesters participation ranged from 5% to 43%, with an overall average participation rate of 16%.

Each group of students was given 1½ hours to complete the instruments. They were provided with paper copies of the questionnaires, a computer answer sheet for the knowledge questionnaire, and a pencil. On the 3 measures, participants were identified by an I.D. number given at the time of administration. These numbers were used to assemble the data for each participant. Only the Participant Key connected the participants' names to the assigned I.D. number.. The consent forms and Participant Key are stored separately from the completed measures in the Research Office and the Principal Investigator does not have access to them. The instruments were returned to the Principal Investigator by the Research Assistant. These steps ensured that participants' names cannot be associated with the collected data. As incentives, a canvas bag imprinted with "Nursing Research is my Bag" or a $10 coupon for Starbucks were provided to participants. A pizza lunch was provided as the sessions occurred during the students' lunch breaks. University of Maryland IRB/Human Research Protocols approval as a minimal risk study was obtained. No unanticipated problems occurred during administration.

Instruments

The research team administered a 90-minute study composed of three instruments: 20 domain knowledge multiple choice questions, 11 interest and activity items, and a written case scenario exercise, based on the maternity nursing domain.

Maternity Nursing Expertise Leveled Questionnaire (ELQ)

The Domain Knowledge multiple choice questions were developed based on a review of the topics covered by five commonly used maternity nursing textbooks. Twenty topics were chosen that were covered by all 5 textbooks and that covered the breadth of the domain of maternal-newborn-women's health nursing. The content of each question was developed to ask about key, central information on the topic. The Cronbach's alpha of this scale of 20 items using the dichotomously keyed correct answers was α=0.851. Previous research with the MDL utilized polytomous scoring in order to increase reliability (Lawless & Kulikowich, 2005). Each knowledge questions had a correct answer and 3 distracters that were categorized not only as wrong but also graded at different levels of expertise in maternity nursing. Three distracters were developed for each question that reflected the range of understanding possible on the topic. For the consumer level incorrect answer, 1 point was given, for the scientist level incorrect, answer 2 points were given, for the competent level answer 3 points were given, and for the proficient correct answer, 4 points were given. An expert panel of three expert nurses reviewed the

instrument for content validity for the correct answer, with an interater agreement of 90%. A sample question is shown in Figure 2.

FIGURE 2. *Example of Expertise-Leveled Question (ELQ).*

Q 1: In fetal circulation:

a. The fetus is protected from environmental toxins by the placenta.
(consumer, incorrect, 1 point given)

b. The umbilical artery carries oxygenated blood from the maternal blood to the fetal superior vena cava.
(scientist , incorrect, 2 points given)

c. The umbilical vein carries deoxygenated blood back to the fetus.
(competent nurse, incorrect, 3 points given)

d. The ductus arteriosus allows the lungs to be mostly bypassed .
(proficient nurse, correct, 4 points given)

For this sample of nursing students, for the 20 questions, the average number of Consumer level answers was 2, of Scientist level answers was 3, of Competent Nurse level answers was 4, and of Proficient Nurse level answers was 11.

To further test the validity of the polytomous scoring within each item, a Pearson correlation was performed. Each question was correlated with the total score on the questionnaire. Ten of the 20 items had correlations with the total score that were significant at the 0.05 or 0.001 level (see Table 2). These 10 items were retained to construct the Maternity Nursing Domain Knowledge Scale. The Cronbach's alpha for the polytomously scored knowledge scale was $\alpha = .349$, compared to 0.851 for the dichotomously scored scale. Previous research with this type of knowledge instrument also indicated slightly less reliability of this type of knowledge scale (Dinsmore, Alexander, & Loughlin, 2008). For this pilot study the polytomously scored Maternity Nursing scale was used in order to maintain comparability to previous MDL research methodology.

TABLE 2. *Maternity Nursing Domain Knowledge Scale: Correlation of polytomously scored variables with Total Score.*

Variables	Pearson *r*	Significance
Q1 Fetal Circulation	0.365**	.001
Q2 Pregnancy Nutrition	0.274**	.010
Q3 Pregnancy Lab Values	0.233*	.025
Q4 Fetal Monitoring	-0.002	.492
Q5 Non-pharmacologic Pain Relief	0.044	.357
Q6 Postpartum Physical Assessment	0.166	.081
Q7 Newborn Metabolic Screening	0.119	.160
Q8 Newborn Physical Assessment	0.339**	.002
Q9 Newborn Jaundice	0.172	.075
Q10 Contraception	0.318**	.003
Q11 Breastfeeding Instruction	0.169	.078
Q12 Pregnancy Screening	0.323**	.003
Q13 High-Risk Pregnancy	0.522**	.000
Q14 Infertility	0.211*	.038
Q15 Menopause	0.197*	.049
Q16 Sexual Development	0.426**	.000
Q17 Reproductive Cancer	0.189	.056
Q18 Breast Conditions	0.141	.119
Q19 Bereavement	0.113	.176
Q20 Professionalism	-0.025	.420

*p<.05 **p<.001

Maternity Nursing Interest Survey

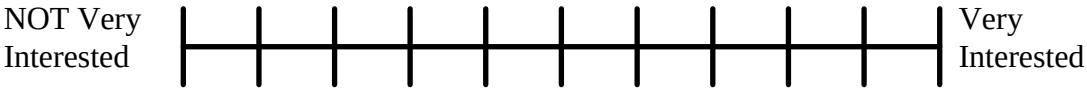
The Maternity Nursing Interest Survey was adapted from other MDL Interest instruments (Dinsmore, Alexander, & Loughlin, 2008). For the five interest questions, 10-cm lines were used to solicit a student's level of interest in maternity nursing topics such as fetal monitoring. The endpoints of the line were identified as *not very interested* (0) and *very interested* (10). If the

student marked an X at the midpoint of the line, a 5 was entered for the variable. Lines were measured with standard rulers, providing interval-level data. Interrater agreement for 20% of the surveys was 100%. An example of an item from this survey is provided in Figure 3.

FIGURE 3. *Sample question from Maternity Nursing Interest Survey.*

For the following items, indicate your interest in the following activities by **marking a line** on the bar that describes your level of interest:

a. Electronic Fetal Monitoring



Scores from these five items were measured and scored on a 1-10 cm scale. These five items were summed to create a Maternity Nursing Interest Survey. Cronbach's alpha for this scale was α=0.851, which was deemed acceptable.

Maternity Nursing Critical Thinking Scenario

Previous research with the MDL measured deep and surface-processing during reading, and other activities. This clinical nursing adaptation, the Maternity Nursing Critical Thinking Scenario (MNCTS, see Figure 4), analyzes a case study that had been extensively used to capture components of critical thinking and clinical reasoning. Students responded to a paper-and-pencil performance task that is typical for nursing. The written clinical scenario provided direct and indirect cues. Students were instructed to list all the patient problems, also known as nursing diagnoses, suggested by the scenario, the priority of each problem, the evidence that led to a patient problem being identified, the important missing data points, relevant nursing interventions, and legal and ethical issues inherent in the case. The participants were also asked to list discharge instructions, however 29% of the students did not provide discharge teaching points, possibly due to the placement of this part of the assignment at the end of the long instrument, so this portion of the instrument was not analyzed. The participants were also asked to list outcome goals, however due to lack of content variability (many students had answers like "stable" or "no complications", so this question was also not analyzed. Inter-rater reliability by two expert maternity nurses for coding of the key used to score the scenarios was 85%.

FIGURE 4. *Text of the Maternity Nursing Critical Thinking Scenario Instrument.*

Critical Thinking in Maternity Nursing

Please consider the following case study:

A.W., an 18 y. o. G2P0010, came to labor and delivery with her boyfriend with a complaint of spontaneous onset of contractions beginning at 1 am. It is now 6 am. She goes to the bathroom to put on a patient gown and to give a urine sample, and stops to breath with a contraction. She says she has had a bit of a headache, for which she took some acetaminophen, and she reports some heartburn. When she returns to bed, she mentions she had recently voided and had a bowel movement. Her membranes have not ruptured. She lies down in bed and you place her on the fetal monitor. The heart tones are heard in the upper right quadrant. You assess the contractions as every 5 minutes and mild to moderate intensity. The fetal heart is 150 bpm with 2-5 bpm variability with the fetal heart going to the 140's after the peak of a contraction. A.W.'s blood pressure is 146/88; her urine sample has +2 protein and trace glucose.

Complete the following questions in relation to THIS case study and use the format of the boxes below the questions. Be sure to put your name on every page. Use as many or as few pages as you need. **Note the last page for discharge planning and family collaboration on page III-7.**

1. What are your priorities in this scenario (Nursing diagnoses, Patient problems)?
2. What evidence is present to support your priorities? How good is the evidence?
3. What else do I need to know? What am I missing?
4. What nursing interventions are appropriate in this situation (based on my priorities and evidence)? In what order should these interventions be implemented?
5. How do I evaluate outcomes in this situation?
6. Are there any legal and/or ethical implications inherent in the scenario or in the nursing interventions I should implement?
7. What is the appropriate discharge planning and collaboration with the family?

This scenario and format was chosen because it is a typical performance activity in nursing education at all levels. The questions correspond to the components of the critical thinking definition described by Scheffer and Rubenfeld (2000) and used as the definition of CT for this study. The steps in the scenario analysis process can also be compared to deep and

surface processing as described in the Model of Domain Learning (Alexander, 2004), and critical thinking researchers (Braten & Olaussen, 2007). The alignment of these variables and components is shown in Table 3.

TABLE 3. *Comparison of MNCTS variables to Critical Thinking Definition components and Strategic Processing components*

Variable Description in Maternity Nursing Critical Thinking Scenario	***Equivalent Components in Critical Thinking definition***	***Equivalent Component Deep or Surface Strategic Processing in Model of Domain Learning***
Identify problems in list	Analyzing	Surface
Didn't identify wrong problems	Discriminating	Surface
Prioritization of problems in correct order	Applying Standards	Surface
Amount of inference required to identify problem based on keyed depth of problem	Logical Reasoning	Deep
Identify cues and evidence to confirm problem	Logical Reasoning	Deep
Identify Missing data needed to care for patient	Information Seeking	Deep
List Interventions needed to care for patient	Transforming Knowledge	Deep
List patient outcome goals.	Predicting	Surface
List Legal and Ethical issues	Predicting	Deep

The critical thinking variables were

1. NUMPROBS, the number of correct patient problems identified by the participant
2. NUMEVIDENCE, the number of correct cues or connections to evidence of patient problems listed in the scenario
3. NUMMISSING, the number of missing data points, salient pieces of knowledge needed to analyze the scenario

4. NUMWRONG, the number of wrong problems the participant listed that were not in the key.

5. NUMINTERVENTIONS, the number of correct nursing actions or interventions the participant listed compared to the key.

6. NUMLEGETH, the number of legal and ethical implications for the patient problem identified by the participant.

7. PRIORITZN, a numerical comparison of the prioritization assigned to all the problems by the participant compared to the keyed prioritization. Each correct problem in the key had a correct prioritization identified. See Appendix A for an explanation of the algorithm used.

8. WTDSUMDEPTH, the weighted sum of the depth of the patient problems identified. Depth refers to the amount of inference required to identify a problem. Each correct problem was rated in the key on depth with a score of 1 to 3, where obvious problems requiring little inference received an 1 and subtle problems requiring a great deal of inference and knowledge of cues received a 3. To derive this variable, the depth scores for the problem that the participant identified were summed.

Results

Maternity Nursing Expertise-Leveled Questionnaire and Maternity Nursing Interest Survey

In order to address the first study objective of examining whether the Model of Domain Learning can be extended to Maternity Nursing, the knowledge and interest scales were examined for differences between groups to see if expected changes occurred. The students in the third semester were compared to fourth semester students that specialized in maternity nursing. An increase in knowledge and interest is generally predicted between acclimation and competence by the MDL, so the students with increased class and clinical time in maternity nursing in the fourth semester would be expected to demonstrate an increase. An independent samples *t*-test was conducted to compare knowledge and interest scale scores for the students in semester 3 to the students in semester 4 who specialized in Maternity Nursing. Results are displayed in Table 4. For the Maternity Nursing Domain Knowledge Scale, there were no significant differences in scores between 3rd semester (*M*=30.60, SD=4.14) and fourth semester (*M*=27.80, SD=4.60; *t* (45)=1.41, p=0.165, two-tailed). For the Maternity Nursing Interest Survey, there was a significant difference in scores between 3rd semester (*M*=32.85, SD=11.55 and fourth semester (*M*=44.62, SD=4.67; *t* (55)= -2.46, p=0.017, two-tailed). The magnitude of the difference in the means (mean difference= -11.77 , 95% *CI:* -21.37 to -2.17) was moderate and statistically significant for the differences in the semesters on the Maternity Nursing Interest Survey (Cohen's d= -1.34, effect size r=0.55.

TABLE 4. *t-Test Results Comparing Knowledge and Interest Scales by Semesters*

Scale	Groups	M	SD	N	*t*-Test	Significance	*df*
Knowledge	3rd Sem	30.60	4.14	42	1.41 [Cohen's d= 0.64 Effect size r= 0.30][1]	.165	45
	4th Sem	27.80	4.60	5			
Interest	3rd Sem	32.85	11.55	51	-2.46 Cohen's d=1.34 Effect size r=0.55	.017*	55
	4th Sem	44.62	4.67	6			

*p<.05

[1]See discussion in Results about reporting effect size with non-significant results

Maternity Nursing Critical Thinking Scenario Analysis

In order to address the second study objective of The responses to the MNCTSA were coded and analyzed. The means and standard deviations for the variables used in the Critical Thinking scale is reported in Table 5.

TABLE 5. *Critical Thinking variables Descriptive Statistics*

VARIABLE	N	MEAN	SD
Number of Problems	85	2.01	1.09
Prioritization of Problems	85	0.73	0.08
Depth of Problems	84	3.79	2.30
Evidence Items	84	1.35	1.07
Missing Data	85	1.87	1.29
Nursing Interventions	85	3.16	2.14
Legal Ethical Implications	85	0.66	0.95
Wrong Problems Listed	85	0.82	0.97

The differences in critical thinking variables between the groups were analyzed in Table 6. For the Maternity Nursing Critical Thinking Scenario analysis, for the variable Correct Evidence listed, there was a significant difference in scores between 3rd semester (*M*=1.10, SD=0.95 and fourth semester (*M*=2.14, SD=1.07; *t* (55)= -2.67, p=0.01, two-tailed). The rest of the critical thinking variables had non significant differences except for the legal ethical implications variable, which had statistically significant results in the non-hypothesized direction: 3rd semester (*M*=0.62, SD=0.90 and fourth semester (*M*=0.14, SD=0.38; *t* (55)= 1.38, p=0.02, two-tailed).

TABLE 6. *t-Test Results Comparing Critical Thinking variables by Semesters*

Variable	Groups	*M*	SD	N	*t*-Test	Significance	*df*
Number of Problems	3rd Sem	2.14	1.07	50	1.00	.32	55
	4th Sem	1.71	0.95	7			
Prioritization	3rd Sem	0.75	0.06	50	-0.41	.68	55
	4th Sem	0.76	0.09	7			
Depth	3rd Sem	4.02	2.35	50	0.48	.64	55
	4th Sem	3.57	2.23	7			
Evidence	3rd Sem	1.10	0.95	50	-2.67	.01*	55
	4th Sem	2.14	1.07	7			
Missing Data	3rd Sem	1.72	1.23	50	-0.54	.59	55
	4th Sem	2.00	1.63	7			
Interventions	3rd Sem	3.46	2.21	50	-0.13	.90	55
	4th Sem	3.57	1.81	7			
Legal Ethical	3rd Sem	0.62	0.90	50	2.49	.02*	55
	4th Sem	0.14	0.38	7			
Wrong Problems	3rd Sem	0.78	0.95	50	0.93	.36	55
	4th Sem	0.43	0.79	7			

*p<.05

To address the second study objective to determine if critical thinking could be objectively measured in a written case scenario in the domain of maternity nursing, and to assist in scale development, a factor analysis was performed. The eight items of the Maternity Nursing Critical Thinking Scenario Analysis were subjected to Principal Components Analysis (PCA) using SPSS version 17. Prior to performing this analysis, the suitability of this data for Factor Analysis was assessed. The ratio of participants to items was greater than ten to one (87:8). Inspection of the correlation matrix revealed the presence of many coefficients of 0.3 and above (see Table 7). The Kaiser-Meyer-Olkin value was 0.59, rounding to meet the recommended value of 0.6 and Bartlett's Test of Sphericity reached statistical significance with a significance value of .00 (Pallant, 2010). All of these factors indicate an adequate level of support for factorability of the correlation matrix.

TABLE 7: *Critical Thinking Variables Correlation Matrix*

	Problems listed	Evidence Items	Missing Data Items	Inter-ventions	Legal Ethical impli-cations	Prioriti-zation of problems	Depth of Problems	Number of Wrong Problems
Problems listed	1.000							
Evidence Items	.150	1.000						
Missing Data Items	**.273**	**.263**	1.000					
Interventions	**.384**	**.293**	**.319**	1.000				
Legal Ethical implications	.097	.194	**.442**	**.334**	1.000			
Prioritization of problems	**.671**	.155	-.029	**.306**	-.078	1.000		
Depth of Problems	**.907**	**.265**	**.268**	**.389**	.109	**.714**	1.000	
Number of Wrong Problems	-.134	-.116	.230	-.026	.142	**-.768**	**-.258**	1.000

Correlations that round up to an acceptable 0.3 are boldfaced.

Principal Components Analysis revealed the presence of three components with eigenvalues exceeding 0.9, explaining 39%, 24%, and 12% of the variance respectively (see Table 8). An inspection of the screeplot revealed a clear elbow break at the third component. To aid in the interpretation of these three components, varimax rotation was performed. The rotated solution revealed the presence of simple structure, with all three components showing a number of strong loadings and all variables loading most substantially on only one component (See Table 9). The interpretation of the first two factors was consistent with previous research on critical thinking with problem identification/surface processing items loading on factor 1 Problem Identification, and problem analysis/deep processing variables loading on factor 2 Problem Analysis. The third factor Problem Specificity had a strong loading for one variable, wrong problem listed.

TABLE 8: *Factor Analysis Eigenvalues and Variance Explained*

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.138	39.226	39.226	2.498	31.223	31.223
2	1.919	23.985	63.211	1.883	23.538	54.761
3	.980	12.247	75.458	1.656	20.697	75.458
4	.738	9.228	84.686			
5	.628	7.849	92.534			
6	.472	5.896	98.430			
7	.087	1.086	99.516			
8	.039	.484	100.000			

TABLE 9: *Varimax Rotation Pattern/Structure Coefficients*

Rotated Component Matrix[a]

	Component		
	1 Problem Identification	2 Problem Analysis	3 Problem Specificity
Correct Problems	.958		
Depth of Problems	.914		
Prioritization of problems	.702		-.673
Legal Ethical implications		.721	
Evidence Items		.712	-.404
Missing Data Items		.657	.373
Interventions	.387	.599	
Number of Wrong Problems			.904

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 7 iterations.

Table 10 shows the mapping of the Critical Thinking scenario variables onto the PCA factors 1, 2 and 3, identified as Problem Identification/Surface Processing, Problem Analysis/Deep Processing, and Problem Specificity/Wrong Problems.

TABLE 10. *Concordances between MNCTS and PCA Factors 1 Problem Identification, 2 Problem Analysis, and 3 Problem Specificity.*

Maternity Nursing Critical Thinking Scenario Variables	**Factors**
Identify problems in list	1: Problem Identification
Prioritization of problems in correct order	1: Problem Identification
Amount of inference required to identify problem based on keyed depth of problem	1: Problem Identification
Identify cues and evidence to confirm problem	2: Problem Analysis
Identify Missing data needed to care for patient	2: Problem Analysis
List Interventions needed to care for patient	2: Problem Analysis
List Legal and Ethical issues	2: Problem Analysis
Didn't identify wrong problems	3: Problem Specificity

Reliability analyses were conducted of the critical thinking scales based on the factor analysis. The Cronbach's alpha for the Problem Identification Scale was α=.65 . The Cronbach's alpha for the Problem Analysis Scale was α=.60. These are borderline acceptable statistics.

Discussion

The objectives guiding this research study were 1) to determine if a theoretical model and instruments used to explain changes in knowledge, interest, and strategic processing in reading and other academic domains could be extended to a clinical domain such as maternity nursing, and 2) to determine if critical thinking could be objectively measured in a written case scenario format in the domain of maternity nursing.

The Model of Domain Learning Applied to Nursing Education

For the first objective, fit of the model to maternity nursing was tested by comparing means for the knowledge and interest to the expected changes predicted by the Model of Domain Learning. The expected changes in knowledge were not confirmed. Possible explanations for this include a small, possibly non-representative sample of nursing students, the low reliability of the polytomously scored Maternity Nursing Domain Knowledge scale compared to the

dichotomously scored scale, and the fact that the two groups were very close in professional development. Although faculty anecdotally report increases in student abilities over college education, the ranges in individual differences may obliterate these differences with cross-sectional data such as used in this study. McMullen and McMullen's study (2008) did find increases using longitudinal data. Cluster analysis would also improve the ability of this research to validate the use of the MDL for education research in nursing.

Another explanation may be the "intermediate effect" noted in other expertise literature. If learners have learned more but have not yet organized that knowledge then the expected increase in learning might not be reflected (Patel, Glaser, & Arocha, 2000). Polytomous scoring is an interesting methodology for assessment and formative feedback to students in nursing. Distractor development and testing may preclude wide application but the scoring can be motivational to students, as partial learning is acknowledged. Instruments need to evolve to meet increasing demands on professionals. The polytomous scoring is a quantitative way of capturing what educators have known for years, that some distractors demonstrate more knowledge and thinking than others.

The interest scale demonstrated a significant, moderate effect size in the predicted direction in semester group differences, explaining 55% of the variance. This well-tested scale is very promising for measuring personal interest in maternity nursing. A sample with a broader range of expertise in the learners would be needed to confirm this scale. A chasm exists between the cognitive and phenomenological approach to the development of expertise in nursing. A connecting factor may be the role of motivation (Field 2004). This scale with an affective component is a positive addition to the study of the development of expertise in maternity nursing.

Many other affective aspects to nursing care such the effect of the nurse's relationship with the patient on clinical outcomes, and the role of the nurse's beliefs in his/her patient care planning would be additions to the understanding of decision-making in nursing. The term critical thinking should evolve into a broader understanding of cognitive, psychomotor, and affective aspects of nursing care. Leading nurse researchers such as Benner and colleagues (2010) and Tanner (2006) are calling for this broader analysis of nursing care also. Overall mixed support for the extension of the MDL to maternity nursing was found, with the Maternity Nursing Interest Scale affirming it's predicted changes, and the Maternity Nursing Domain Knowledge Scale demonstrating the need for further validation to be useful.

Measurement of Critical Thinking

To address the second objective to determine if critical thinking could be objectively measured in a written case scenario format in the domain of maternity nursing, a typical nursing written performance was elicited from the participants. Written case scenarios have drawbacks since they are static and do not reflect internal processes. For this instrument, is it a step backwards to have a written scenario? (c.f. Ericsson p.6). Kamin and colleagues' critical thinking instrument analysis (2000) comparing text case descriptions to video descriptions found that the text cases did a good job of capturing aspects of CT, so I felt it was tenable to use a written case

scenario format. Theoretically (Table 3), the CT variables mapped well onto the CT definition used in this study from the Consensus process, and onto the MDL deep and surface processing aspects. However, only one variable actually performed in the expected manner with statistical significance: Evidence items listed by the participant. This is an interesting finding because being able to link assessment and history data to possible patient problems is a key nursing action. I was surprised that more analyses of CT variables were not significant, but the very small N for the fourth semester maternity nursing specialty group made it difficult to achieve significance even if a true difference existed. Larger sample size in the future could overcome this limitation.

Two other challenges encountered in previous literature were constraints here also. In spite of incentives, the "exit phenomena" may have compromised an accurate reading of the description of fourth semester students as they charge toward graduation and dismiss testing. Recruitment also was a strong challenge for a detailed research instrument that is real "work" and not just a survey of attitudes or a self-report evaluation. Greater mentoring of student nurses into research culture will assist with this, as will the increased emphasis on doctorally prepared faculty and the push for evidence-based teaching as well as evidence-based clinical practice, so that participation in research is a valued and expected activity.

As discussed in the review of the literature, previous studies using teacher-made tests or instruments were often based on definitions of critical thinking without a model of how the variables were related. This pilot study showed somewhat promising results by using a well-tested model of the learner development process.

Another challenge to generalization is that many CT instruments are embedded in teaching strategies, so that a broader understanding of the development of clinical thinking in nursing cannot be identified since the instruments cannot be used with practicing nurses. Practicing nurses as well as acclimating students in must be considered when developing items and instruments.

The factor analysis produced some very promising confirmations of congruence between the CT variables in the MNCTS and previous CT definitions. The relatively close mapping of the PCA factors onto the MDL component Strategic Processing, surface and deep aspects, contributes to a more favorable evaluation of the ability to measure CT with a written case study. The reliability of the Problem Identification Scale and Problem Analysis Scale also provided moderate indirect support for the coherence of these scales for future use.

Implications for Future Research

More testing and refinement of the Maternity Nursing Domain Knowledge Scale should be done to increase reliability and validity. A greater quantity of items are needed, and more rigorous validity testing needs to occur. The Interest Survey should be administered to participants with a broader range of expertise. Improvements to the critical thinking scenario

include the migration of the survey to an online environment. This could also allow for unfolding scenarios to be presented and more precision in understanding the use of cues.

An important methodology that would contribute to an improvement in the Critical Thinking Scenario analysis is think-alouds. With a wide range of participant nurses at the acclimating, competent, and proficient level, insight on the process of critical thinking could lead to a better scale. These types of improvements will require greater funding for nursing education research. Advocacy by all nurses for increased federal nursing education funding may contribute to more resources becoming available.

The challenge of quality assurance discussed in the introduction of this paper could be addressed by using the patient outcomes measures as outcome benchmarks. Some outcomes that could be measured include length of stay, patient satisfaction or pain scale, cost of care measures, efficiency of care with time stamps, quality and quantity of interaction among disciplines.

One problem identified in the literature that was not well addressed by this study is control of moderators. Factors such as GPA, age, race, type of prerequisite education could have an influence on this process and they are not well understood.

The research presented here offers a domain-specific, quantitative, replicable methodology to analyze the development of CT in nurses across their professional growth. The Model of Domain Learning provided a framework that guided analysis and reflected current understandings of CT in nursing literature. An interest scale with good reliability was adapted and predictive validity in future study. This study provided a few baby steps forward, but much qualitative and quantitative research to build on the science of measuring nursing expertise development remains to be done. Theories, instruments, and methodologies such as those suggested by Model of Domain Learning research are promising resources for this journey.

REFERENCE LIST

Abel, W.M., & Freeze M. (2006) Evaluation of concept mapping in an associate degree nursing program. *Journal of Nursing Education, 45(9),* 356-364.

Abrami, P.C., Bernard. R.M., Borokhovski, E., Wade, A., Surkes, M.A., Tamim, R., & Zhang, D. (2008.) Instructional interventions affecting critical thinking skills and dispositions: a stage 1 meta-analysis. *Review of Educational Research, 78*(4), 1102-1134

Alexander, P.A. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational, and strategic forces. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 213-250). Greenwich, CT: JAI Press.

Alexander, P. A., Sperl, C. T., Buehl, M. M., Fives, H., & Chiu, S. (2004a). Modeling domain learning: Profiles from the field of special education. *Journal of Educational Psychology, 96*(3), 545-557.

Alexander P.A. (2004) A model of domain learning: reinterpreting expertise as a multidimensional, multistage process. In: *Motivation, Emotion, and Cognition: Integrative Perspectives on Intellectual Functioning and Development* (eds D.Y. Dai & R.J. Sternberg). Erlbaum, Mahwah, New Jersey.

Alfaro-LeFevre, R. (1999). *Critical Thinking in Nursing: A Practical Approach,* (2nd ed.) Philadelphia: WB Saunders Co.

American Association of Colleges of Nursing (April 13, 2010) Patient Protection and Affordable Care Act. Retrieved from http://www.aacn.nche.edu/Government/pdf/HCRreview.pdf.

American Association of Colleges of Nursing (October 20, 2008).The essentials of baccalaureate education for professional nursing practice. Retrieved from http://www.aacn.nche.edu/education/pdf/BaccEssentials08.pdf

Beckie, T. M., Lowry, L.W., & Barnett, S. (2001). Assessing critical thinking in baccalaureate nursing students: A longitudinal study. *Holistic nursing practice, 15*, 18-26.

Benner, P.E., Hughes, R.G.& Sutphen, M. (2008). Clinical Reasoning, Decisionmaking, and Action: Thinking Critically and Clinically. In Hughes, R.G. (ed.). Patient safety and quality: An evidence-based handbook for nurses. (Prepared with support from the Robert Wood Johnson Foundation). AHRQ Publication No. 08-0043. Rockville, MD: Agency for Healthcare Research and Quality; March 2008.

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

Braten, I., & Olaussen, B.S. (2007). The motivational development of Norwegian nursing students over the college years. *Learning in Health and Social Care, 6(1),* 27-43.

Brunt, B. A. (2005). Critical thinking in nursing: An integrated review. *Journal of Continuing Education in Nursing, 36*(2), 60-67.

Cohen, S.B., Grote, K. D., Pietraszek, W. E., & Laflamme, F. (2010). Increasing consumerism in healthcare through intelligent information technology. *American Journal of Managed Care, 16,* SP37-43.

Daly, W.M. (1998). Critical thinking as an outcome of nursing education. What is it? Why is it important to nursing practice? *Journal of Advanced Nursing, 28,* 323-331.

Daly, W. M. (2001). The development of an alternative method in the assessment of critical thinking as an outcome of nursing education. *Journal of Advanced Nursing, 36(1),* 120-130.

Dinsmore, D. H., Alexander, P. A., & Loughlin, S. M. (2008). The impact of new learning environments in an engineering design course. *Instructional Science, 36,* 375-393.

Edwards, S. (2006). Critical thing at the bedside: a practical perspective. *British Journal of Nursing,12(19),* 1142-1149.

Ennis, R.H.(1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher, 18(3),* 4-10.

Ennis, R.H. & Weir, E. (1985). The Ennis-Weir critical thinking essay test. Pacific Grove, CA: Midwest Publications.

Facione, P.A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Research findings and recommendations.* Newark, DE: American Philosophical Association.

Facione, N.C.,& Facione, P.A. (1994) the California Critical Thinking Skills Test and the National League for Nursing Accreditation Requirement in Critical Thinking. California Academic Press

Field, D. E. (2004). Moving from novice to expert—The value of learning in clinical practice: A literature review. *Nurse Education Today, 24,* 560-565.

Garcia, T., Pintrich, P.R. (1992). Critical thinking and its relationship to motivation, learning strategies, and classroom experience, *Paper presented at the Annual Meeting of the American Psychological Association* (100th, Washington, D.C., August 14-18, 1992).

Hoffman, J.J. (2006). *The relationships between critical thinking, program outcomes, and NCLEX-RN performance in traditional and accelerated nursing students.* Unpublished doctoral dissertation, University of Maryland School of Nursing, Baltimore, MD.

Hsu, L-L. (2004). Developing concept maps from problem based learning scenario discussions. *Journal of Advanced Nursing, 48(5),* 510-518.

Institute of Medicine. (2010). *The future of nursing: Leading change, advancing health.* Washington, D.C.: National Academies Press.

Itano, J.K. (1989). A comparison of the clinical judgment process in experienced registered nurses and student nurses. *Journal of Nursing Education, 28(3),* 120-126.

Jacobsen, L. A. (2011). America's Aging Population. *Population Bulletin 66 (1).*

Jennings, B.M. (2008). Patient acuity. In Hughes, R. G. (Ed.). *Patient safety and quality: An evidence-based handbook for nurses.*

Jones, S. & Brown, L. (1991). Critical thinking: impact on nursing education. *Journal of Advanced Nursing, 16,* 529-533.

Kamin, C. S., O'Sullivan, P. S., Younger, M., & Deterding, R. (2001). Measuring critical thinking in problem-based learning discourse. *Teaching and Learning in Medicine, 13(1),* 27-35.

Kataoka-Yahiro, M. & Saylor, C. (1994). A critical thinking model for nursing judgment. *Journal of Nursing Education, 33(8),* 351-356.

Lasater, K, (2007) Clinical judgment development: Using simulation to create an assessment rubric. *Journal of Nursing Education 46(11)* 496-503.

Lasater, K, & Nielsen, A. (2009) The influence of concept-based learning activities on students' clinical judgment development. *Journal of Nursing Education 48(8)* 441-446.

Lawless, K. A., & Kulikowich, J. M. (2005). Domain knowledge and individual interest: The effects of academic level and specialization in statistics and psychology. *Cotemporary Educational Psychology, 31,* 30-43.

McMullen, W.F., & McMullen, C. (2008). Evaluation, inference, and analysis: Patterns of change in critical thinking skills of graduate nursing students. *Paper presented at annual meeting of the American Educational Research Association* (March 25-27, New York, NY).

Oermann, M.H. (1997). Evaluating critical thinking in clinical practice. *Nurse Educator, 22(5),* 25-28.

Oermann, M.H. &Gaberson, K. (1998). *Evaluation and Testing in Nursing Education*. New York, Springer.

Pallant, J. (2010). *SPSS Survival Manual (4th Ed).* New York: McGraw Hill.

Patel, V.L., Glaser, R., Arocha, J.F. (2000). Cognition and Expertise: Acquisition of Medical Competence. *Clinical and Investigative Medicine, 23(4,;* 256-60.

Pintrich, P.R.,. Smith, D.A.F., Garcia, T., McKeachie, W.J., (1991). A m*anual for the use of the motivated strategies for learning questionnaire (MSLQ).* Ann Arbor, MI: NCRIPTAL, The University of Michigan.

Robert Wood Johnson Foundation (December, 2008). Reducing the administrative burden of health care quality reporting (Vol XI, No. 10). Retrieved from http://www.academyhealth.org/files/pdf/findings1208.pdf

Roberts, J.D. (2000). Problem-solving skills fo senior student nurses: an exploratory study using simulation. *International Journal of Nursing Studies, 37*, 135-143.

Samuelstuen, M. S. & Braten, I. (2007). Examining the validity of self-reports on scales measuring students' strategic processing. *British Journal of Educational Psychology, 77*, 351-378.

Scheffer, B., & Rubenfeld, M. G. (2000). A consensus statement on critical thinking in nursing. *Journal of Nursing Education*, *39*, 352-30. 195-6.

Simon, H. (2001). Problem-solving. In R. A. Wilson & F. C. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences* (pp. 674-676). Cambridge, MA: The MIT Press.

Simpson, E., & Courtney, M. D. (2002). Critical thinking in nursing education: A literature review. *International Journal of Nursing Practice*, *8*(April), 89-98.

Tanner, C.A. (1983). Research on clinical judgment In W. Holze*mer (Ed), Review of research in nursing education* (pp. 1-32). Thorofare, N.J. : Slack.

Tanner, C. A. (1997). Spock would have been a terrible nurse (and other issues related to critical thinking in nursing). *Journal of Nursing Education (36(1)* 3-4.

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

Tishman, S. & Andrade, A. (N.D.). Thinking dispositions: A review of current theories, practices, and issues. Retrieved April 26, 2009 from http://learnweb.harvard.edu/alps/thinking/docs/dispositons.htm

Walsh, C.M., & Seldomridge, L.A., (2006a). Measuring critical thinking: One step forward, one step back. *Nurse Educator, 31(4),* 15-162.

Walsh, C.M., & Seldomridge, L.A., (2006b). Critical thinking: back to square two. *Journal of Nursing Education, 45(6),* 212-219.

Walters, J.S. (1986). Critical thinking in liberal education: a case of overkill? *Liberal Education, 72*, 233-244.

Watson, G., & Glaser, E.M. (1980). *Watson-Glaser critical thinking appraisal manual*. Cleveland: Psychological Corporation.

Weinstein, C.E., Husman, J. & Vierking, D.R., (2000). Self-regulation interventions with a focus on learning strategies. In: M. Boekaerts, P.R Pintrich, & M. Zeidner (Eds), *Handbook of self-regulation* (pp 727-745) San Diego, CA: Academic Press.

APPENDIX A.

Assigning a Score to an Open-Ended List of Priorities

Introduction

When a nurse examines a patient, the nurse must identify the problems that the patient is experiencing, assign a priority to the problems and treat the most urgent problems first. This ability is a critical thinking skill that student nurses need to learn during their education. Nursing instructors can evaluate this skill in students by presenting them with a scenario in which information about a patient is presented and having the student write down the problems the patient is experiencing and the priority or each problem.

The nursing instructor can compare the ability of different students by assigning scores to the set of responses given by the students. The purpose of this section is to a method for assigning such scores. This method was created by Richard B. Winston and Lily Fountain.

Methodology

To assign a score, the instructor must first generate a key in which all the problems are identified and assigned priorities. The priorities must be positive integers with 1 being the highest priority. The numbers need not be consecutive and ties are allowed. The priority assigned to each item should reflect the severity of the consequences for the patient if the item is missed. Thus, if the severities of two items are similar, those items should be assigned similar priorities. Conversely, if the severities of two items are dissimilar, those items should be assigned dissimilar priorities. The instructor must also designate a priority code for incorrect responses by the students and an artificial code which marks the end of the responses by the students. The wrong response code and artificial code must be different from any of the priority codes assigned to any of the correct responses. The assignment of priorities is a subjective process but once the priorities are assigned, the remainder of this method is objective.

To score an individual student, the instructor first lists the correct priority of each item that the student identified in the priority order used by the student. For example, suppose the problems in the key were

- not breathing
- unconscious
- minor abrasions
- homelessness

The code for wrong responses in this example is 97 and the artificial code is -1. To the list of items in the key, is added a code for wrong response. The final list for the key would be as follows.

KEY
1, 2, 3, 4, 97

In this example, the problems listed by the student in order were

- not breathing
- bruises
- minor abrasions.

The instructor would make the following list

1, 97, 3

Next the instructor adds the artificial code to the end of the list and then adds all the items that the student did not identify in reverse priority order. The list would now be

List for student
1, 97, 3, -1, 4, 2

The next step is to identify how far each item in the list is from the beginning of the list for the key after having removed all previous items except the code for wrong answers from the list for the key. The artificial code is skipped when assigning scores to each item. The distance for any item after the artificial code has been encountered is increased by one. The sum of all those distances is a measure of how poor a student's list of priorities is with higher scores representing a poorer performance. Generally, it is more convenient for a high score to represent a good performance rather than a poor one and to scale the score from zero to 1. To achieve this, the student's score is subtracted from the highest possible score and then divided by the highest possible score. The highest possible score is calculated using an artificial priority list in which all the responses are wrong and the number of responses is equal to the maximum number of items identified by any student. (With the key listed above and a maximum number of responses by any student equal to 5, the maximum possible score is 26.)

The scores for individual items would be assigned as follows:

- For item 1 (1), the score is zero because item 1 is at the beginning of the list for the key. Item 1 is removed from the list for the key and the modified list for the key is now 2, 3, 4, 97
- For item 2 (97), the score is 3 because item 2 is the last item in the list and must be moved 3 positions to become the first item. The list for the key is not modified because the code for wrong answers is never removed from the list.

- For item 3 (3), the score is 1 because it must be moved 1 positions to become the first item . Item 3 is removed from the list for the key and the modified list for the key is now 2, 4, 97
- Item 4 (-1) represents the end of the students responses. It is skipped.
- For item 5 (4), the score is 2 because it must be moved 1 positions to become the first item and then the distance is increased by 1 because the student never listed this item as a priority. Item 5 is removed from the list for the key and the modified list for the key is now 2, 97
- For item 6 (2), the score is 1 because item 1 is at the beginning of the list for the key giving it a distance of zero and then the distance is increased by 1 because the student never listed this item as a priority. Item 1 is removed from the list for the key and the modified list for the key is now 97

The steps in the score procedure are listed in the table below.

Step	**KEY**	**List for student**	**Score**	**Explanation**
1	**1**, 2, 3, 4, 97	**1**, 97, 3, -1, 4, 2	0	"1" doesn't have to be moved to get to the beginning of the key.
2	2, 3, 4, **97**	1, **97**, 3, -1, 4, 2	3	"97" must be moved 3 spaces to get to the beginning of the key.
3	2, **3**, 4, 97	1, 97, **3**, -1, 4, 2	1	"3" must be moved 1 space to get to the beginning of the key.
4	2, 4, 97	1, 97, 3, **-1**, 4, 2	0 (skipped)	-1, the code for the end of the answers given by the student is skipped.
5	2, **4**, 97	1, 97, 3, -1, **4**, 2	2	"4" must be moved 1 space to get to the beginning of the key. A penalty of 1 is added because the student didn't give this answer.
6	**2**, 97	1, 97, 3, -1, 4, **2**	1	"2" doesn't have to be moved to get to the beginning of the key but 1 penalty of 1 is added because the student

				didn’t give this answer.

The total score 0 + 3 + 1 + 2 + 1 = 7. This is subtracted from the maximum possible score (26) and scaled to attain the final score of (26 – 7)/26 = 0.73.

Discussion

The method described above provides a consistent method for assigning scores to open-ended lists of prioritized items. It assigns higher scores to prioritized lists in which more items were correctly identified correctly and also for assigning the correct priorities among the items that were identified. It does not address two issues. (1) Students who make no responses can get a better score than students who make some correct and some incorrect responses. (2) The scores of all the students depend of the maximum number of answers made by any student because the key must be at least as long as the number of answers by any student. As a case in point, in the example above, if another student had identified not breathing as the first priority and then had given four wrong answers, the key would end up as 1, 2, 3, 4, 97, and the score for the first student would be 0.77 instead of 0.73.

AUTHOR NOTE:
I would like to thank the University of Maryland School of Nursing Dean's Teaching Scholar Grant, which funded part of this research. I am especially grateful to my Alexander lab colleagues Emily Fox, Daniel Dinsmore, Liliana Maggioni, Sandra Loughlin, Alexandra List, Meghan Parkinson, Emily Grossnickle, and Peter Bagetta, as well as my advisor Patricia Alexander, and my nursing colleagues/mentors Janice Hoffman and Jeanne Geiger-Brown. My wonderful husband Richard Winston provided extensive assistance with manuscript preparation. Many thanks to all!