Development and Psychometric Testing of the Relational Inquiry Capacities Scale

by © Ahtisham Younas

A practicum report submitted to the

School of Graduate Studies

in partial fulfillment of the requirements for the degree of

Master of Nursing

School of Nursing

Memorial University of Newfoundland

August 2017

St John's Newfoundland and Labrador

Abstract

Background: One of the major challenges of nursing educators is adequately preparing students to nurse effectively and efficiently within a demanding health care system. Relational Inquiry, a new nursing practice approach, prompts nursing students to embrace the complexity through reflecting on a patient's situation at intrapersonal, interpersonal, and contextual levels. However, students must be equipped with several capacities such as "compassion" and "curiosity," to practice from this approach yet there are no instruments to assess these necessary capacities. The Relational Inquiry Capacities Scale (RICS) was developed to address this gap.

Purpose: To develop the RICS and to determine the scale's psychometric properties.

Methods: The RICS was developed using Relational Inquiry as the guiding framework; an integrated literature review of existing scales that measure any of the relational capacities; a modified Delphi technique of expert consultations with Relational Inquiry scholars; and, assessment of the scale's readability and comprehension by one linguist, two laypersons, and five nursing students. Psychometric testing was completed through a pilot study by distributing the scale to a sample of nursing students from three nursing schools.

Findings: Based on the Relational Inquiry approach, an integrative literature review, and the modified Delphi technique, the RICS item pool was revised resulting in 73 items under six subscales: compassion, self-compassion, curious, competence, commitment and correspondence. Psychometric testing indicated that nine items should be excluded from the RICS.

Conclusion: The RICS is a good measurement instrument and it is recommended that it be employed by nursing educators to assess the extent to which nursing students have acquired the relational capacities necessary for quality nursing care delivery.

Acknowledgements

I would like to thank and acknowledge those individuals who supported me during this Master of Nursing Degree program.

Special thanks to my practicum supervisor, Dr. Caroline Porr, Associate Professor, Memorial University of Newfoundland School of Nursing for her continuous support, guidance, and encouragement throughout this research practicum. I also would like to extend a special thanks to Ms. Maria Kelsey and Father Wayne Bolton S.J., for proof reading my assignments during this Master's degree and for being available to me whenever I needed support and assistance. I would like to thank Father Earl Smith S.J. for his support and encouragement. To all those who participated in the consultation process, thank you for your feedback and suggestions which made this practicum possible.

Finally, I would like to pay a tribute to my mother, Josephine Younas, and my eldest brother Robert, who worked hard and made countless sacrifices to help me achieve my career goals, and encouraged, supported and believed in me. You taught me how to be patient and determined and how to achieve my goals with the limited resources available. To my brother, Albert, thank you for encouraging me to pursue higher education from outside my home country and for guiding me at every step of this master's degree. To my brothers Patras, Baber, and Waqas, I thank you for your support and encouragement and the many sacrifices you made for me. To my sister Sana, thank you for all your love and support during this academic journey. Without all of you, I would not have achieved this accomplishment.

Table of Contents

Development and Psychometric Testing of the Relational Inquiry Capacities Scale

Hartrick-Doane and Varcoe (2015) proposed that their Relational Inquiry approach can serve as a guide for practice for both nurses and nursing students. This approach was developed within the Canadian context and entails two components: "relational consciousness" and "inquiry as a form of action" (p. 3). Relational consciousness refers to the expectation that the nurse's role entails deliberately examining relational experiences of patients, families, and other people involved in any given situation in order to understand different factors influencing the clinical situation. Inquiry as a form of action refers to the significance of critically examining each nursing situation and understanding all the influential factors and perspectives of the persons involved (Hartrick-Doane & Varcoe, 2015).

In order to practice from the Relational Inquiry perspective, nurses need to expand their thinking; that is, they need to reflect on patient situations and the context, which the co-developers have categorized as layers of interpersonal, intrapersonal, and contextual factors that shape nurse-patient interactions (Hartrick-Doane, 2014; Hartrick-Doane & Varcoe, 2015). The co-developers explained that the ability to intervene and provide effective nursing care depends on relational capacities: "compassion," "self-compassion," "curiosity," "commitment," "competence," and "correspondence" (Hartrick-Doane & Varcoe, 2015). The Relational Inquiry approach has been purported to be an effective guide for nursing educators for bridging the gap between nursing theory and practice thereby helping students integrate theoretical knowledge into practice (Hartrick-Doane & Varcoe, 2015; Hartrick-Doane,2002a). This approach can also serve

as a means to assess the capacity of students to provide quality nursing care (Hartrick-Doane & Varcoe, 2015). Nursing educators, then, require ways to assess students' appropriation of the six aforementioned relational capacities. The purpose of this research practicum was to develop such an instrument (i.e., RICS) for nursing educators in order that they can advance student abilities to provide high quality nursing care in increasingly complex and challenging health care environments.

Significance of Relational Inquiry

Relational Inquiry based teaching could be described as a process for assisting nursing students to translate theoretical nursing knowledge, skills, values, and ideologies into effective nursing actions (Hartrick-Doane & Varcoe, 2015). The effectiveness of the Relational Inquiry approach has not been studied extensively in nursing education research. However, several authors have considered this approach as a guide for nursing educators who want to revisit their teaching and learning pedagogies. Hartrick-Doane (2002a) argued that behavioral teaching pedagogies such as traditional lectures, skill performance, and so forth, can only assist nursing students to learn technical skills because such pedagogies only cover the course content. Behavioral teaching pedagogies do little to assist students to develop capacities for building meaningful relationships among themselves and with their patients. Therefore, it is recommended that nursing educators should use Relational Inquiry as an educational tool given there is empirical evidence that Relational Inquiry enables nursing students to gain in-depth understanding of nursing concepts (Hartrick-Doane, 2002a).

Consistent with this notion, Hartrick-Doane (2002b) indicated that behavioral teaching pedagogies cannot prepare creative and caring nurses because such pedagogies only help educators to develop technological and empirical nursing skills. Hartrick-Doane moved beyond technical skills when teaching nursing ethics, for example, by engaging students in activities which draw their attention to morally complex and uncertain nursing situations. She was able to foster student creativity and caring abilities and prepare students for ethical nursing practice. Similarly, Spadoni, Doane, Sevean, and Poole (2015), adapted the Relational Inquiry approach to develop a 6-week educational project comprised of mask making and storytelling. The purpose of their project was to help students understand the essence of caring in nursing practice. Through participation in this project the students recognized the importance of caring and connecting with self and others. They were able to improve their self-reflection and awareness skills. They realized that in order to become a caring individual they should be honest about their thoughts about themselves and each other, practice active listening, and analyze their assumptions from various different perspectives.

Rationale for the Research Practicum Project

The previously discussed study and claims by Hartrick-Doane highlight the importance of the Relational Inquiry approach in guiding development of effective teaching and learning strategies. It can serve as an important educational tool. There is little research examining the use of the Relational Inquiry approach in nursing practice and whether nursing students are able to adopt the approach for clinical practice. However, given that the Relational Inquiry approach is promising as a guide for nursing

practice and advancing student learning, a discussion follows exploring what the co-developers of Relational Inquiry describe as key relational nursing capacities (compassion, competence, commitment, correspondence, and curiosity).

The relational capacities, with the addition of *self-compassion*, are essential for enabling nursing students to critically assess any given clinical situation; examine factors influencing the situation; and then, discern their nursing obligations and responsive actions (Hartrick-Doane & Varcoe, 2015). Researchers across the world have highlighted the importance of four capacities namely *compassion, self-compassion, commitment, and competence* (Burnell, & Agan, 2013; Lee & Seomun, 2016a, 2016b; Papadopoulos & Ali, 2016; Yang & Jiang, 2014). However, there is a dearth of research on *curiosity* and *correspondence*. Also missing in the literature, is research substantiating to what degree nurses and nursing students possess these capacities and to what extent these capacities have been operationalized as an interrelated whole. This gap may exist due to limited instruments measuring any of the six capacities (Burnell, & Agan, 2013; Lee & Seomun, 2016b; Papadopoulos, Shea, Taylor, Pezzella, & Foley, 2016; Yang & Jiang, 2014). Therefore, this research practicum began to address the lack of instruments through the development and psychometric testing of a new tool (i.e., the RICS), and, focusing on assessment of nursing students.

Conceptual Definitions of the Relational Capacities

Conceptual definitions of the capacities required to practice from the Relational Inquiry approach and as outlined in *How to Nurse: Relational Inquiry with Individuals and Families in Changing Health and Health Care Contexts*, are outlined below. While

the co-developers presented five relational capacities (i.e., compassion, curiosity, commitment, competence, and correspondence), I decided to integrate self-compassion with compassion. Within the RICS I measured self-compassion as a separate capacity.

Compassion

Compassion means "to share suffering, being in solidarity with persons, and doing with one another" (Hartrick-Doane & Varcoe, 2015, p. 103-104).

Self-Compassion

Self-compassion is the process of "consciously and intentionally choosing to act and respond to yourself and the situation to promote well-being at interpersonal, intrapersonal, and contextual levels" (Hartrick-Doane & Varcoe, 2015, p.113).

Curious

Being curious is about "being interested, inquisitive, and open to uncertainty that is part of disease and illness experiences. It is the capacity to work between knowing and not knowing" (Hartrick-Doane & Varcoe, 2015, p.115).

Commitment

"Being committed means to actively and consciously identify the values and concerns that orient" one's work as a nurse and "continually monitor" how one's actions are "aligning with those commitments" (Hartrick-Doane & Varcoe, 2015, p.118).

Competence

Being competent is "not just about oneself and the knowledge and skills one possesses. Rather, it is person/context dependent and is determined in and by what transpires in particular relational situations" (Hartrick-Doane & Varcoe, 2015, p.125).

Correspondence

Correspondence refers to "relating to and with people in a way that is meaningful to them. It involves paying attention to the meaning, concerns, and life situations of people and families" (Hartrick-Doane & Varcoe, 2015, p.129).

Research Practicum Purpose and Objectives

The overall purpose of this research practicum was to construct the RICS for nursing educators to measure six relational capacities in nursing students that are key to practicing from a Relational Inquiry approach and then to determine the extent to which the new scale is valid and reliable (i.e., the psychometric properties). To achieve this purpose, I pursued the following objectives:

1. To differentiate the aims and format of a questionnaire and an inventory, from those of a scale.
2. To describe and then apply the steps of scale development for survey research.
3. To understand the role and processes of a modified Delphi technique and expert consultation in scale development.
4. To utilize existing scales and nursing theories to inform scale development.
5. To learn the process of determining the psychometric properties of a newly developed assessment scale.
6. To demonstrate advanced nursing practice competencies in the domains of research and consultation and collaboration.

Research Practicum Methods

Four methods were used for the development and psychometric testing of the RICS. First, a methodology paper about scale development was written to fully outline and explain the process for developing a new scale to assess student relational capacities and how to determine the scale's psychometric properties. Second, an integrative literature review was conducted to identify existing questionnaires and scales that measure the relational capacities to inform the development of the RICS item pool. Third, a modified Delphi technique involving 11 Relational Inquiry scholars was implemented to evaluate the relevance of the RICS (i.e., the item pool) with nursing practice and with *How to Nurse: Relational Inquiry with Individuals and Families in Changing Health and Health Care Contexts*. In addition, one linguist, two laypersons, and five nursing students assessed its readability and comprehension. Fourth, psychometric testing of the RICS was completed through a pilot study of nursing students enrolled at three nursing institutions.

Scale Development Methodology Paper

I was able to understand how to develop an assessment scale by exploring and writing about the process of scale development in a paper that was later submitted for publication. I also included the steps for determining the psychometric properties. Psychometric properties refer to a scale's statistical measurement strengths and weaknesses. Psychometric properties provide researchers with important information about how well a scale can measure the intended constructs of interest (DeVellis, 2016).

I was also able to differentiate between three interchangeable terms used in the literature—a scale, a questionnaire, and an inventory. The scale is the most standardized

and precise instrument of data collection compared to a questionnaire and an inventory (DeCoster, 2005; Grove, Burns, & Gray, 2013). Questionnaires generally measure knowledge, beliefs, opinions, or perspectives of participants. Scales generally measure attributes or dimensions of different attributes of participants on a continuum (Grove et al., 2013). In contrast to questionnaires and scales, inventories are catalogues of different attributes, attitudes, perceptions, beliefs, and so forth. Inventories are used to examine certain characteristics or traits of participants ("Tools of Research," n.d.).

Within this methodology paper I delineated a 6-step process for scale development. The steps were selection, conceptualization, and contextualization of constructs; item generation; scaling of items; item selection; psychometric testing; and finalization of the scale. A brief summary of these steps is provided below. The scale methodology paper can be found in Appendix I.

I explained that a scale can measure a single construct or multiple constructs (or variables). These constructs are conceptualized based on type, dimensions, and attributes. Broadly, constructs are grouped into observable, behavioural or cognitive categories (DeVellis, 2016; Waltz, Strickland, & Lenz, 2010). In order to conceptualize a scale's constructs researchers should have knowledge of the type, dimensions, and attributes of each construct (Polit & Beck, 2014). This knowledge can be gained through in-depth study, literature reviews, theoretical and conceptual frameworks, and concept analyses. The RICS is a multiple construct instrument designed to measure behavioral and cognitive constructs (i.e., nursing capacities including compassion, self-compassion,

curiosity, commitment, competence, and correspondence). These six capacities were conceptualized based chiefly on the Relational Inquiry approach.

Several methods for item generation are outlined in this methodology paper including literature reviews, focus group and individual interviews, the Delphi technique, expert consultations, and, existing theories, models, and conceptual frameworks. The methods used for item generation of the RICS were: existing theory (i.e., the Relational Inquiry approach); an integrative literature review; and, the modified Delphi Technique using expert consultations. From among the three standard methods for item scaling—direct, comparative, and econometric estimation—the RICS items were scaled using the direct estimation methods (i.e., the forced Likert Scale consisting of strongly agree= 6, agree= 5, slightly agree=4, slightly disagree=3, disagree= 2, and strongly disagree= 1).

I also explained that a scale's psychometric properties should be assessed through pilot studies and then large-scale validity and reliability studies. Pilot studies are designed to perform preliminary psychometric testing of the developed scale (Thabane et al., 2010) with indicators that include several types of reliability (test-retest, inter-rater, parallel form, and internal consistency) as well as several types of validity (face, content, construct, and criterion). The psychometric properties of the newly developed RICS were assessed through a pilot study of undergraduate nursing students using the measures of: internal consistency (i.e., Cronbach's alpha and Guttmann lambda-2 [λ-2]), total item correlation, inter-item correlation, face validity, and content validity.

Integrative Literature Review

The integrative literature review was conducted to develop an item pool for the RICS. The literature search was performed within the databases: CINHAL, PubMed, Science Direct, and Google Scholar. Based on this literature review the conceptual and operational definitions of the scale's constructs (i.e., the relational capacities) were compared with the conceptual definitions proposed in the Relational Inquiry approach. The items of the questionnaires and scales identified in the literature were critically analyzed and the items found consistent with Relational Inquiry were included in the item pool for the RICS. The major findings of the review were:

1. Out of 80 eligible studies, 15 studies were selected for review, of which 6 focused on the development and psychometric testing of scales to measure compassion (Burnell & Agan, 2013; Lee & Seomun, 2016; Lown, Muncer, & Chadwick, 2015; Neff, 2003; Pommier, 2010; Raes, Pommier, Neff, & Van Gucht, 2011) and 8 outlined the process for developing nursing competence scales (Cowan, Wilson-Barnett, Norman, & Murrells, 2008; Hsu & Hsieh, 2013; Lin, Hsu, Li, Mathers, & Huang, 2010; Liu, Kunaiktikul, Senaratana, Tonmukayakul, & Eriksen, 2007; Meretoja, Isoaho, & Leino-Kilpi, 2004; Nilsson et al., 2014; Safadi, Jaradeh, Bandak, & Froelicher, 2010; Takase & Teraoka, 2011). One study focused on measuring nurses' commitment (Lin, Wang, Li, & Huang, 2007); however, the full text article and scale used could not be located. One study focused on the development of a scale to measure curiosity (Kashdan et al., 2009). There were no studies on scale development for measuring nurses' correspondence.

2. Few scales existed that measure relational capacities. Only scales measuring compassion, competence, and curiosity were identified. Commitment and correspondence scales located in the review were not pertinent to nursing students; rather, they measured career satisfaction and choices, professional commitment, professional development, and organizational commitment of nurses.
3. The comparisons of the identified scales with the extant definitions of the relational capacities indicated that some of the scales were consistent and others were somewhat consistent with the Relational Inquiry approach. However, the majority of the scale items were inconsistent. Therefore, the RICS was developed by including selected items from the reviewed scales and remaining items were developed from the assumptions of the Relational Inquiry approach.

In total, 21 items were extracted and adapted from the reviewed scales and then included in the item pool for the RICS. The remaining items were drawn from the assumptions of relational capacities according to the Relational Inquiry approach. Finally, a preliminary draft of the RICS was developed which comprised 6 subscales and 72 items. The six subscales consisted of: compassion (15 items), self-compassion (13 items), curious (9 items), commitment (10 items), competence (16 items), and correspondence (9 items). The integrative literature review for developing the item pool of the RICS can be found in Appendix II.

The Modified Delphi Technique and Expert Consultations

The primary purpose of the modified Delphi technique and expert consultations was to finalize the items for the RICS and to ensure face validity. Consultations were conducted on an ongoing basis throughout the research practicum. First, during the phases of proposal development and scale methodology paper writing, I consulted two statisticians (Dr. Yanqing Yi and Dr. Veeresh Gadag) from Memorial University of Newfoundland School of Medicine. These experts provided suggestions concerning the process for: developing an item pool; scaling the items; scoring the items; as well as how to test for validity and reliability. Second, during conceptualization and operationalization of the RICS, Dr. Hartrick-Doane's guidance was sought in order to clarify both philosophical and practical assumptions of the requisite relational capacities and to identify potential Relational Inquiry experts across Canada for reviewing the RICS. Third, based on the suggestions by Dr. Hartrick-Doane, the first draft of the RICS was sent to other Relational Inquiry experts including Relational Inquiry co-developer, Dr. Varcoe.

Initially, the item pool comprised 90 items before a review by five undergraduates and graduate nursing students to assess consistency of items with daily nursing practice. Based on their review comments, the item pool was reduced to 72 items. This 72-item pool was then sent for review to the co-developers of the Relational Inquiry approach and three other Relational Inquiry scholars in Canada. Based on their review and suggestions, several items were re-worded and one item was added. The revised item pool of 73 items was again sent to the nursing students for evaluating the readability and

comprehensibility of the items. In addition, at this stage, two laypersons and one linguistic were consulted to evaluate the readability and comprehensibility of the items. Finally, the revised item pool was then sent to another eight Relational Inquiry scholars and nursing educators at Memorial University of Newfoundland School of Nursing. Six scholars agreed to review the scale for its consistency with nursing practice. Based on this final review, the items were re-worded again and the RICS was finalized for pilot testing.

Psychometric Testing of the Relational Inquiry Capacities Scale

Study Design and Setting

A pilot study of the RICS was conducted in the School of Nursing, Memorial University of Newfoundland, at the Center for Nursing Studies, and at the Western Regional School of Nursing after ethical clearance was sought from the Interdisciplinary Committee on Ethics in Human Research and participating schools of nursing. As aforementioned, the purpose of the study was to assess the reliability and validity of the RICS. Please see Appendix IV for ethics approval letter.

Target Population and Data Collection

The target population comprised all the undergraduate nursing students enrolled at the nursing institutions. Consistent with convenience sampling, all the students were invited to participate in the study and those agreeing to participate completed an online survey. Therefore, no specific inclusion and exclusion criteria were established. The RICS was uploaded on Survey Monkey and a flyer detailing study information, the link to the online survey, and contact information of the researchers was sent to potential

study participants through the program administrator of the undergraduate nursing program of the respective schools of nursing. Survey responses were collected between August 1st and August 15th of this year.

Data Analysis

Completed surveys were checked for errors; then coded and the data entered into Statistical Package for the Social Sciences (SPSS 24.0) for analysis. Descriptive analysis was performed including measures of central tendency and dispersion of the values for each of RCIS items. Reliability testing was performed by calculating Cronbach's alpha and λ-2, and inter-item-correlation. Validity testing, comprising face and content validity, was performed during the RICS development phase (i.e., the modified Delphi technique and expert consultations).

Ethical Considerations

Strict protocols were set to maintain privacy and confidentiality of participants. Data access was only granted to the researchers and data files were password protected. The computers that stored the data were also password protected. Identifying information including names, electronic mail addresses, and student numbers was not collected. The Internet Protocol addresses of participating students who completed the online survey were not visible to the researchers.

Results

In total, 18 students responded to the pilot survey. Out of the 18 surveys completed by participants, 14 surveys were fully completed, and hence were included in the analysis. The incomplete surveys were excluded from the analysis. In general, the

competence subscale was found to be the most reliable subscale and the curious subscale was the least reliable subscale. Reliability testing results are discussed below.

Internal Consistency

The internal consistency of the six subscaleswere assessed using three measures—Cronbach's alpha, λ-2, and inter-item-correlation. The interpretation of reliability measures was based on the rules outlined below.

1. The acceptable value of Cronbach's alpha ranges from 0.70 to 0.95 (Bajpai & Bajpai, 2014; Streiner et al., 2015; Tavakol & Dennick, 2011).
2. The acceptable item-total correlation of each item with the total scale should be more than 0.20 (Streiner et al., 2015).
3. When comparing Cronbach's alpha value of each item with the item-total correlation, the item-total correlation value should not be lower than 0.20. If lower, compare the "Alpha If Item Deleted" value with the overall Cronbach's alpha of the RCIS. If the "Alpha If Item Deleted" value of the item is greater than the Cronbach's alpha of the RCIS, then delete the item (Oppenheim, 1992).
4. If the Cronbach's alpha value of each subscale is less than 0.70 compare it with the λ-2. The λ-2 statistic assesses the variance in data due to true scores. If respondents differ significantly in their abilities to answer the scale items, λ-2 will be high and the error will be low. The higher the λ-2, the better the scale is able to differentiate between respondents' abilities. If λ-2 is greater than 0.80, the Cronbach's alpha value may not be robust and its low value is merely the property of data (Callender & Osburn, 1979; Osburn, 2000).

5. If Cronbach's alpha and λ-2 values are less than 0.80, assess the "Alpha If Item Deleted" value. This value indicates the change in Cronbach's alpha if the item is removed from the scale. If an item removed leads to improvements in the RCIS's alpha then the item could be removed (Callender & Osburn, 1979; Osburn, 2000).

Compassion. The Cronbach's alpha of the compassion subscale was 0.74 and λ-2 was 0.78 whereas the alpha for the items ranged from 0.70 to 0.79. Out of 16 items, 3 items had low item-total correlation and Alpha If Item Deleted Values included: *I do not impose my own ideas when caring for patients* (item-total correlation = 0.021 and Alpha If Item Deleted= 0.76); *I do not try to fix patients' emotional problems* (item-total correlation = 0.036 and Alpha If Item Deleted= 0.79), and *I respect the dignity of patients* (item-total correlation = -0.079 and Alpha If Item Deleted= 0.76). Since both alpha and λ-2 values were less than 0.80, these three items were excluded. The post-exclusion analysis showed an increased alpha (0.83) and λ-2 (0.84).

Self-Compassion. The Cronbach's alpha of the compassion subscale was 0.79 and λ-2 was 0.83 whereas the alpha for the items ranged from 0.76 to 0.81. Out of 13 items, 1 item had low item-total correlation and Alpha If Item Deleted Values included: *I often reflect on how I care for patients* (item-total correlation = 0.0053 and Alpha If Item Deleted= 0.81). Since alpha of the total subscale was lower than λ-2, the Alpha If Item Deleted Value of the item was compared with the total alpha of the RCIS. The item was then excluded. The post-exclusion analysis showed an increased alpha (0.81) and λ-2 (0.84).

Curious. The Cronbach's alpha of the compassion subscale was 0.62 and λ-2 was 0.70 whereas the alpha for the items ranged from 0.54 to 0.66. Out of 9 items, 1 item had low item-total correlation and Alpha If Item Deleted Values included: *I consciously examine my knowledge about patients' needs* (item-total correlation = -0.025 and Alpha If Item Deleted= 0.66). Since alpha of the total subscale was lower than λ-2, the Alpha If Item Deleted Value of the item was compared with the total alpha of the RCIS. The item was then excluded. The post-exclusion analysis showed an increased alpha (0.66), but the λ-2 decreased (0.65). Since both of these values were equal and statistically non-significant, it indicated that the curious subscale actually possesses low internal consistency.

Commitment. The Cronbach's alpha of the compassion subscale was 0.68 and λ-2 was 0.72 whereas the alpha for the items ranged from 0.58 to 0.73. Out of 10 items, 2 items had low item-total correlation and Alpha If Item Deleted Values included: *I do not overburden myself while caring for patients* (item-total correlation = 0.14 and Alpha If Item Deleted= 0.73) and *I try to fulfill my commitments with patients* (item-total correlation = 0.092 and Alpha If Item Deleted= 0.70). Since alpha of the total subscale was lower than λ-2, the Alpha If Item Deleted Value of the item was compared with the total alpha of the RCIS. The items were then excluded. The post-exclusion analysis showed an increased alpha (0.74) and λ-2 (0.77).

Competence. The Cronbach's alpha of the compassion subscale was 0.88 and λ-2 was 0.90 whereas the alpha for the items ranged from 0.86 to 0.89. Out of 16 items, 2 items had low item-total correlation and high Alpha If Item Deleted Values including: *I*

consciously seek answers to any questions arising from my nursing practice (item-total correlation = 0.13 and Alpha If Item Deleted= 0.88) and *I prioritize my nursing actions according to changing situations* (item-total correlation = 0.13 and Alpha If Item Deleted= 0.88). Since both alpha λ-2 were higher and equal, the items with low item-total correlation were excluded. The post-exclusion analysis showed an increased alpha (0.89) and λ-2 (0.91).

Correspondence. The Cronbach's alpha of the compassion subscale was 0.72 and λ-2 was 0.76 whereas the alpha for the items ranged from 0.64 to 0.73. None of the items in this subscale had low item-total correlation values. The alpha and λ-2 values were approximately equal.

Face Validity

In terms of face validity of the RCIS, Relational Inquiry scholars and nursing students indicated that the RICS is consistent with the nursing practice approach in their respective settings and that the RICS can be used to measure the six relational capacities.

Content Validity

Nine Relational Inquiry scholars reviewed the RICS for congruence with Relational Inquiry. They agreed that the RICS can measure the six relational capacities. From among the nine scholars, two scholars indicated that some items should be included to better represent contextual factors relevant to each of the relational capacities. Since seven out of nine scholars judged the RICS highly relevant, the content validity index was given a value of 78%. All nursing students, questioned, indicated that the RICS can measure the six relational capacities and that respective items are consistent with

everyday nursing practice. The valuations indicate that the RICS has acceptable content validity.

Pilot Study Conclusions

The pilot study enabled me to conclude that the RICS is a good instrument for assessing relational capacities of nursing students and I recommend the RICS for use in nursing curriculum; and additionally, as a tool for nursing researchers wanting to conduct survey research. Moreover, before excluding the subscale items with low alpha value, in the future I plan to carry out further testing using a larger sample to reassess construct validity and to reevaluate reliability.

Advanced Nursing Practice Competencies

The Canadian Nurses Association (2008) outlined several essential nursing competencies for advanced practice that require graduate-level education. Competencies fall under four domains: clinical, research, leadership, and, consultation and collaboration. During this research practicum project, I gained and demonstrated advanced knowledge, skills, judgment and personal attributes under the domains of research, and, consultation and collaboration.

Research Competencies

Regarding research competencies, the Canadian Nurses Association (2008) advises that an advanced practice nurse should be able to “work as primary investigator or collaborator to conduct research for the benefit of nursing practice” (p. 23). Consistent with this competency, I worked as a primary investigator to develop an item pool for the RICS. I first completed a scale development methodology paper and an integrative

literature review which enabled me to enhance my knowledge of scale development. By conducting the pilot study to determine the RCIS psychometric properties, I advanced my skills of scale development, and at the same time, demonstrated my ability to assume the role and responsibilities of a principal investigator. Other expectations of an advanced practice nurse under this domain are the abilities to "critique, interpret, apply, and disseminate evidence-based findings" (p. 24). I critiqued empirical studies and analyzed available scales on six nursing capacities (compassion, self-compassion, competence, curiosity, commitment, and correspondence). I also analyzed items in existing scales through the lens of the Relational Inquiry approach thereby enhancing my critical thinking skills. Another expectation is the ability to "disseminate new information through formal and informal channels, including presentation and publication, at local, regional, national, and international levels" (p. 24). Consistent with this, I presented the process and results of this research practicum at Memorial University of Newfoundland School of Nursing and submitted a manuscript for publication to an international peer-reviewed journal.

Consultation and Collaboration Competencies

In terms of the consultation and collaboration competencies, an advanced practice nurse should collaborate with other nurses and healthcare team members in a timely manner to promote quality improvement (CNA, 2008). Consistent with this, I engaged in lengthy discussions (in person and through communication technologies) with various nursing and non-nursing experts to develop an item pool for the RICS which enabled me to improve my consultation skills. The Canadian Nurses Association (2008) also

recommends that nurses should learn to “practice collaboratively and build effective coalitions” (p. 26). Consistent with this, when I discussed my research practicum project with the co-developers of the Relational Inquiry approach I demonstrated my ability to receive constructive feedback and accept alternative ideas or recommendations. I also sought advice for the project from other Relational Inquiry scholars, some of whom may not have agreed with my ideas. I was prepared to modify the RCIS accordingly. I also engaged in consultation and collaboration with members of the Interdisciplinary Committee on Ethics in Human Research during proposal writing and the ethics approval phase of the research practicum project.

Concluding Remarks

This research practicum project led to the development and pilot testing of the RICS to assess the levels of six relational capacities: compassion, self-compassion, curiosity, commitment, competence, and correspondence, in nursing students. It is anticipated the RICS will assist nursing educators to assess and evaluate the extent to which their students possess these essential capacities to nurse in complex healthcare environments, and, serve as the basis for remedial learning. Based on this assessment, educators will be able revisit classroom and clinical teaching pedagogies and strategies to assist students to further enhancing these capacities.

References

Bajpai, S., & Bajpai, R. (2014). Goodness of measurement: Reliability and validity. *International Journal of Medical Science and Public Health, 3* (2), 112-115. doi: 10.5455/ijmsph.2013.191120133

Burnell, L., & Agan, D. L. (2013). Compassionate care: Can it be defined and measured? The development of the compassionate care assessment tool. *International Journal of Caring Sciences*, *6*(2), 180-187.

Callender, J. C., & Osburn, H. G. (1979). An empirical comparison of Coefficient alpha, Guttmann's lambda-2, and MSPLIT maximized split-half reliability estimates. *Journal of Educational Measurement*, *16*(2), 89-99.

Canadian Nurses Association (CNA). (2008). Advanced nursing practice a national framework. Retrieved from: https://www.cna-aiic.ca/~/media/cna/page-content/pdf-en/anp_national_framework_e.pdf

Cowan, D. T., Wilson-Barnett, D. J., Norman, I. J., & Murrells, T. (2008). Measuring nursing competence: Development of a self-assessment tool for general nurses across Europe. *International Journal of Nursing Studies*, *45*(6), 902-913.

DeCoster, J. (2005). Scale construction notes. [Online]. Retrieved from http://www.stat-help.com/scale_20100316.pdf

Grove, S. K., Burns, N., & Gray, J. (2013). *The practice of nursing research: Appraisal, synthesis, and generation of evidence* (7th ed.). St. Louis, MO. Elsevier.

Hartrick-Doane, G. (2002a). Beyond behavioral skills to human involved processes: Relational nursing practice and interpretive pedagogy. *Journal of Nursing Education*, *41*(9), 400-404.

Hartrick-Doane, G. (2002b). In the spirit of creativity: The learning and teaching of ethics in nursing. *Journal of Advanced Nursing*, *39*(6), 521-528.

Hartrick-Doane, G. (2014). Beyond relational oblivion. *Advances in Nursing Science*, *37*(2), 84-86.

Hartrick-Doane, G., & Varcoe, C. (2015). *How to nurse: Relational inquiry with individuals and families in changing health and health care contexts*. Philadelphia, PA: Lippincott.

Hsu, L. L., & Hsieh, S. I. (2013). Development and psychometric evaluation of the competency inventory for nursing students: A learning outcome perspective. *Nurse Education Today*, *33*(5), 492-497.

Kashdan, T. B., Gallagher, M. W., Silvia, P. J., Winterstein, B. P., Breen, W. E., Terhar, D., & Steger, M. F. (2009). The curiosity and exploration inventory-II: Development, factor structure, and psychometrics. *Journal of Research in Personality*, *43*(6), 987-998.

Lee, Y., & Seomun, G. (2016a). Compassion competence in nurses. *Advances in Nursing Science*, *39*(2), E54-E66.

Lee, Y., & Seomun, G. (2016b). Development and validation of an instrument to measure nurses' compassion competence. *Applied Nursing Research*, *30*, 76-82.

Lin, C. J., Hsu, C. H., Li, T. C., Mathers, N., & Huang, Y. C. (2010). Measuring professional competency of public health nurses: Development of a scale and psychometric evaluation. *Journal of Clinical Nursing*, *19*(21-22), 3161-3170.

Lin, C. J., Wang, H. C., Li, T. C., & Huang, L. C. (2007). Reliability and validity of nurses' job satisfaction scale and nurses' professional commitment. *Mid-Taiwan Journal of Medicine*, 12, 65-75.

Liu, M., Kunaiktikul, W., Senaratana, W., Tonmukayakul, O., & Eriksen, L. (2007). Development of competency inventory for registered nurses in the People's Republic of China: Scale development. *International Journal of Nursing Studies*, *44*(5), 805-813.

Lown, B. A., Muncer, S. J., & Chadwick, R. (2015). Can compassionate healthcare be measured? The Schwartz Center compassionate care scale. *Patient Education and Counseling*, *98*(8), 1005-1010. doi: 10.1016/j.pec.2015.03.019

Meretoja, R., Isoaho, H., & Leino-Kilpi, H. (2004). Nurse competence scale: Development and psychometric testing. *Journal of Advanced Nursing*, *47*(2), 124-133.

Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity*, *2*(3), 223-250.

Nilsson, J., Johansson, E., Egmar, A. C., Florin, J., Leksell, J., Lepp, M., ... & Carlsson, M. (2014). Development and validation of a new tool measuring nurses self-reported professional competence—The Nurse Professional Competence (NPC) Scale. *Nurse Education Today*, *34*(4), 574-580.

Oppenheim, A. N. (1992). Questionnaire design, interviewing and attitude measurement. London: Pinter.

Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods*, *5*(3), 343.Retrieved 3/9/2017 from: https://pdfs.semanticscholar.org/c65a/b8997cb48561a76c922ec971eb434986e78c.pdf

Papadopoulos, I., & Ali, S. (2016). Measuring compassion in nurses and other healthcare professionals: An integrative review. *Nurse Education in Practice*, *16*(1), 133-139.

Papadopoulos, I., Shea, S., Taylor, G., Pezzella, A., & Foley, L. (2016). Developing tools to promote culturally competent compassion, courage, and intercultural communication in healthcare. *Journal of Compassionate Health Care*, *3*(1), 1-10. doi: 10.1186/s40639-016-0019-6

Pommier, E. A. (2010). The compassion scale. Dissertation Abstracts International Section A: Humanities and Social Sciences, 72, 1174.

Raes, F., Pommier, E., Neff, K. D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the self-compassion scale. *Clinical Psychology & Psychotherapy*, *18*(3), 250-255.

Rosenau, P., Watson, L., Vye-Rogers, L., & Dobbs, M. (2015). Educating for complexity in nursing practice: A baccalaureate curriculum innovation. *Quality Advancement in Nursing Education*, *1*(3), 1-14.

Safadi, R., Jaradeh, M., Bandak, A., & Froelicher, E. (2010). Competence assessment of nursing graduates of Jordanian universities. *Nursing & Health Sciences*, *12*(2), 147-154.

Spadoni, M., Doane, G. H., Sevean, P., & Poole, K. (2015). First-year nursing students—developing relational caring practice through inquiry. *Journal of Nursing Education*, *54*(5), 270-275.

Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use*. Oxford, UK: Oxford University Press.

Takase, M., & Teraoka, S. (2011). Development of the holistic nursing competence scale. *Nursing & Health Sciences*, *13*(4), 396-403.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53-55. doi: 10.5116/ijme.4dfb.8dfd

Tools of research. (n.d.). Retrieved from http://results.mu.ac.in/myweb_test/Research%20Methadology-Paper-3/Chapter-9.pdf

Yang, G. F., & Jiang, X. Y. (2014). Self-directed learning readiness and nursing competency among undergraduate nursing students in Fujian province of China. *International Journal of Nursing Sciences*, *1*(3), 255-259.

Appendix I

Scale Development Methodology Paper

Scale Development for Survey Research

Questionnaires and scales are considered the simplest and the most widely used methods for data collection in survey research (Timmins, 2015; Rattray & Jones, 2007). Questionnaires are distinct from scales because the former are less precise and standardized the later (Grove, Burns, & Gray, 2013). A questionnaire is "a spoken, written, or printed form used in gathering information on some participant or participants, consisting of a set of questions that assesses a phenomenon for research purposes" (Holosko and Thyer 2011, p. 99). A scale is defined as "a scheme, inventory, rating form, or device by which some property, attribute, or behavior may be universally measured" (Holosko and Thyer, 2011, p. 111). In recent years, new questionnaires and scales have been developed. However, despite the development and extensive use of questionnaires and scales, nursing literature provides only limited guidance for developing effective questionnaires and scales. Though questionnaires and scales are perceived as simple methods of data collection, they are constructed through a rigorous and systematic process and need much careful work. Some researchers have provided comprehensive guidelines concerning questionnaire development, but the literature pertinent to scale development is limited. For example, Timmins (2015) discussed the function and content of questionnaires; their use in survey research; how to ensure their validity and reliability; and how surveys and questionnaires are developed. Since Timmins's discussion is primarily centered on questionnaire development and survey research, missing is the process of scale development. Likewise, Kumar (2016) explained the process to ensure the validity and reliability of questionnaires and scales but provided minimal discussion

about the process of item generation, scaling, and scale refinement. Also, Kumar's research is based on old literature (from 1992-2007) and did not include recent guidelines about scale development. The aim of this methodology paper is to outline a systematic and rigorous process of scale development for survey research. To clearly explicate the scale development process, an exemplar of how one would develop the Relational Inquiry Capacities Scale (RICS) is presented.

Differentiation between a Questionnaire, a Scale and an Inventory

Prior to discussing the scale development process, further differentiation between the questionnaire and scale is in order, along with the other commonly interchangeably used term in survey research: the inventory. It is important to differentiate between a questionnaire, a scale and an inventory because using these terms interchangeably can be misleading and may create confusion for novice researchers. For example, there are an ample data collection instruments for measuring stress levels of nursing students which include Perceived Stress Scale, Adolescent Coping Orientation for Problem Experiences Inventory, Physio-Psycho-Social Response Scale, Stress and Coping Inventory and Coping Behaviour Inventory and so forth. Such a variety instruments for measuring stress is legitimate because of its abstract nature and differences in opinions about what stress entails, but the inconsistent use of the terms: scale, questionnaire and inventory could leave the researchers confused or mistaken about the format.

There are several similarities and differences between questionnaires and scales. The main similarities are that both can be self-administered or interviewer-administered and both contain two parts, the question or item and a response set. In addition, both

questionnaires and scales are developed through a systematic rigorous process. The four main differences between questionnaires and scales are as follows:

1. Questionnaires and scales differ in levels of precision and standardization. Questionnaires, although broad, are less precise and standardized than scales (Grove, et al., 2013).
2. The response set in questionnaires could be categorical such as “yes” and “no”, but the response set in scales is generally designed on a semantic, visual, Likert, or Guttmann scale (Grove et al., 2013; DeCoster, 2005) for example, a 5-point Likert scale “strongly agree”, “agree”, “neutral”, “disagree”, “strongly disagree”.
3. Questionnaires generally measure knowledge, beliefs, opinions, or perspectives of participants (Grove, Burns, & Gray, 2013). They are used to collect data for “measuring knowledge attitudes beliefs and feelings” (LoBiondo-Wood and Haber 2014, p. 279). For example, a questionnaire may measure the beliefs and perspectives of nurses’ patients regarding compassionate care. On the other hand, scales generally measure attributes or dimensions of different attributes of participants on a continuum (Grove et al., 2013). For example, a scale may measure the levels of compassion of nurses with items entailing different attributes of compassion and a response set that includes “strongly agree”, “agree”, “neutral”, “disagree”, and “strongly disagree”. A researcher may measure the pain of patients, after compassionate care, on a visual analog scale ranging from 0 to 10 (0= no pain 10= severe pain). Questionnaires cannot be used to measure any variable on a continuum such as compassion or competence as

this task is associated with scales. Scales ask respondents to rank some trait or ability on continuum of possible responses" (Houser 2016, p. 198). Therefore, it could be considered that scales are subsets of questionnaires.

4. Questionnaires and scales can measure one or more constructs. However, scales have a moderately large number of items to measure the intended construct with each item possessing a theoretically similar meaning (DeCoster, 2005). For example, a researcher can use negative and positive worded items to measure levels of nurse compassion. Each positive and negative worded item will measure the compassion from a different perspective and the total score will be summed. In questionnaires both positive and negative worded items can be used, but the responses cannot be summed for a total score. It is important to note that "when multiple items are used to measure a single concept, such as quality of life or anxiety, and the scores on those items are combined mathematically to obtain an overall score, the questionnaire or measurement instrument is called a scale" (LoBiondo-Wood & Haber, 2014; p. 280).

In contrast to questionnaires and scales, inventories are catalogues of different attributes, attitudes, perceptions, and beliefs and so forth. Inventories are used to examine certain characteristics or traits of participants. For example, a researcher who wants to measure emotional intelligence of nurses can develop a list of different attributes of emotional intelligence with a response set of "agree" and "disagree" or "yes" or "no". Like questionnaires, inventories could be designed with a categorical response set of yes or no ("Tools of research", n.d.).

Scale Development

Step I: Selection, Conceptualization, and Contextualization of Construct

The first and foremost step of scale development is to outline the purpose of the scale and to indicate the intended constructs, concepts, or variables to be measured (Giesen, Meertens, Vis-Visschers, & Beukenhorst, 2012). Some scales measure a single construct while others measure multiple constructs such as attitudes, capacities, behaviors, and emotions. Therefore, it is important for researchers to clearly establish whether a scale will measure a single construct or multiple constructs (Polit & Beck, 2014), because failure to do so would affect the validity and reliability of the scale. For example, in order to develop RICS, the purpose and the constructs of this scale should be made explicit. The purpose of RICS is to determine levels of six nursing capacities: compassion, curiosity, commitment, competence, and correspondence in nursing students. The RICS will measure 6 capacities which indicates that it will be a multiple construct scale.

Once the construct is selected, the next step is conceptualization of the construct which depends on construct's type, dimensions, and attributes. There are different types of constructs which are broadly categorized into observable or behavioural categories (DeVellis, 2016; Waltz, Strickland, & Lenz, 2010). For example, skin colour and body type are observable whereas patient satisfaction, compassion, and curiosity are behavioural constructs. In order to conceptualize the intended constructs researchers should become knowledgeable of the type, dimensions, and attributes of a construct (Polit & Beck, 2014). This knowledge can be gained through in-depth study, literature reviews,

theoretical and conceptual frameworks, and following concept analyses. Upon gaining a thorough understanding and knowledge, researchers should develop a conceptual definition of the concept and differentiate it from other similar constructs (Furr, 2011). For example, if a researcher intends to measure self-efficacy it is important to differentiate it from self-confidence. Such conceptualization and differentiation from related constructs is critical for the item selection, validity, and reliability of the scale. In relational to the RICS, the six constructs, which are behavioural in nature, will be conceptualize using a nursing philosophy called the "Relational Inquiry" approach proposed by Hartrick-Doane and Varcoe (2015).

Prior to embarking on scale construction, researchers should contextualize the scale by selecting the setting and target population for which the construct is likely to be measured (Polit & Beck, 2014). It is important to outline the features of the setting, the demographics, and the characteristics of the target population (Furr, 2011). For example, a researcher who intends to measure compassion should identify the setting which could be clinical, educational, or community. The population could be practicing nurses, nursing students, educators or nursing managers. Contextualization is necessary to ensure that easily understandable, appropriate, and culturally acceptable items are selected to measure the construct (Polit & Beck, 2014; Grove et al., 2013). During contextualization, it is also imperative to consider the administration method for the scale with respect to the target population, setting, and time frame (Furr, 2011). An inappropriate method of scale administration could pose a cognitive and emotional burden on participants as well as may introduce different kinds of bias such as recall and social desirability bias

(Bowling, 2005). In relation to the RICS, the study population will be undergraduate nursing students studying at Canadian universities and the RICS will be a self-administered scale.

Step II: Item Generation

The second step of scale development is item generation**.** The purpose of item generation is to develop an item pool to operationalize the conceptual definition of the construct that has been established in the first step. This operationalization should be consistent with the scale purpose, target population's characteristics, and adopted conceptual model or framework (Streiner, Norman, & Cairney, 2015; Radhakrishna, 2007). For example, if a researcher intends to develop an instrument to measure compassion levels of nurses and doctors, the generated items should be consistent with the established purpose of the scale because failure to do so could lead to an inaccurate operationalization. Similarly, the items for operationalizing compassion in a nursing and medical context will differ due to the differences in how compassion is perceived in both populations. Also, the generated items should be consistent with the selected nursing and medical model or conceptual framework adopted to conceptualize the construct. Regarding the context of the RICS, the conceptual definitions of the 6 constructs should be truly grounded in the Relational Inquiry approach and the generated items should be consistent with the practical understanding of the six constructs in the nursing profession.

Several methods of item selection have been proposed which can be used without any logical order. Some of commonly used methods are: literature reviews, focus groups and individual interviews, the Delphi Technique, expert consultation, and use of existing

theories, models or conceptual frameworks (Streiner et al., 2015; Waltz et al., 2010). Each of these methods will be briefly discussed.

Literature reviews. Literature reviews serve two main purposes in scale development. First, as previously mentioned, it helps in conceptualization of the construct. Second, it enables researchers to find any existing scales measuring the construct and enables development of an item pool (Artino Jr, La Rochelle, Dezee, & Gehlbach, 2014). The best method of conducting a literature review for item selection is reviewing, analysing, comparing, and critiquing the existing scales and the conceptual definition of the construct under consideration with the researcher's established conceptual definition of the construct (Waltz et al., 2010). For example, if a researcher intends to conduct a literature review to identify existing scales measuring compassion of nurses, he or she should compare the conceptual and operational definitions of compassion located in the identified scales with his or her established conceptual definition. During the literature review, if researchers find an existing scale that is consistent with their own established conceptual definition, a new scale should not be developed. It is important because several scales measuring a similar construct could create confusion. Some researchers may find it daunting to select the appropriate scale for their own research and others may embark on the development of a new scale. In both cases, researchers may miss an opportunity to further validate the existing scales and dismiss them without reasonable justification (Streiner et al., 2015).

Focus groups and individual interviews. Another method of item selection is conducting focus groups or interviews with the key informants representing the target

population. This method is mostly used when the construct to be measured is rarely addressed in the literature. For example, nurses' reflexivity, presencing with patients, and therapeutic use of self could be examples of such concepts which require further analysis (Hessel, 2009). These focus groups and interviews are more focused than those conducted for qualitative research because the participants are expected to suggest broad themes or items pertinent to the intended construct. It the responsibility of the researcher to analyse the broad themes, develop specific items, and then validate the developed items with the participants (Waltz et al., 2010). The main purpose of focus groups in item generation phase is to develop an appropriate question or item list to be included in the constructed scale. These focus groups are exploratory in nature and are not meant to generalize the findings to a specific group or context (Nassar-McMillan and Borders 2002, Masadeh 2012). Hence, these focus groups and interviews are more focused than those conducted for qualitative research because the participants are expected to suggest broad themes or items pertinent to the intended construct. It is the responsibility of researcher to analyse the broad themes, develop specific items, and then validate the developed items with the participants (Waltz *et al* 2010). Streiner et al., (2015) suggested that for item pool development (a list of generated items for operationalizing any particular construct) a maximum of two or three focus groups are adequate with 8-10 people.

The Delphi Technique. The Delphi Technique is a kind of survey method which is used to explore a rarely addressed phenomenon. It can be categorized into conventional Delphi approaches and modified Delphi approach. The conventional Delphi approaches

include, “a Policy Delphi to devise a strategy to address a specific problem; a “Classical Delphi to forecast the future; and, a “Decision-Making Delphi to achieve better decision making” (Avella 2016, p. 306). In conventional approaches, an expert panel respond to open ended questions concerning the phenomenon of interest in several rounds and the researcher analyses their responses to reach a consensus (Avella 2016, Wilkes 2015). However, in the modified Delphi technique a researcher develops “an initial list of responses based on a review of the relevant literature and disseminates it to the expert panel. The panel would then be asked to rank the list according to a specific criterion provided by the researcher” (Avella 2016, p. 313). A panel of experts are selected who express their opinions, thoughts, and judgements about the phenomenon. For example, if a researcher intends to develop a scale to measure preparedness of nursing students for nursing theory-based practice, a group of nursing theorists and nursing educators who teach nursing theories can be included in the Delphi panel. The responsibility of the researcher is to collect and analyse the responses of the panellists, maintain objectivity of the findings, and reach a consensus (Grove et al., 2013; Waltz et al., 2010). Although the use of the Delphi Technique in scale development is not widely discussed in the literature, many researchers believe that it can be used to develop a scale or its item pool (Grove et al., 2013; Waltz et al., 2010).

Expert consultation. The process of expert consultation in scale development is somewhat similar to the Delphi Technique. However, as previously noted, unlike the conventional Delphi approaches, during expert consultation a researcher develops a preliminary item pool, or scale, and sends it for review to the experts (Streiner et al.,

2015). Hence, expert consultation could be considered as a modified Delphi technique (Avella 2016). Based on this expert review, the scale and its items are finalized. This method is also used to ensure the content validity of a scale (Kumar, 2016; Bolarinwa, 2015; Polit & Beck, 2014) which will be discussed later in the paper.

Theories, models or conceptual frameworks. Existing theories and models are commonly used for scale development in nursing, psychology, and social sciences (Grove et al., 2013; Waltz et al., 2010; Furr, 2011). The use of theories enables researchers to outline the dimensions and linkages of the construct under consideration and other related factors (Kumar, 2016). A researcher can select a theory or model which describes the construct to be measured and can generate an item pool based on its assumptions. For example, if a researcher wants to measure caring levels of nursing students several nursing theories such as Watson's (1979) caring model, Leininger's (1993) culture care theory, and Swanson's (1991) caring theory can be adopted to develop a scale or generate an item pool.

All the afore-mentioned methods can be used for item generation. Therefore, it is the researcher's choice to select a method that is feasible and convenient. Another consideration in item generation is how many items are required to adequately operationalize a construct (Artino et al., 2014). Most of the authors have suggested that the number of items depends on the type, dimensions, and complexity of a construct (Streiner et al., 2015; Artino et al., 2014). It has also been suggested that the number of items should be twice the number of items required in the final scale (Grove et al., 2013). Therefore, researchers should generate a large item pool because several items are

usually deleted during validity and reliability testing (Streiner et al., 2015; Waltz et al., 2010). Researchers must also ensure that the generated items are stated in an unambiguous and linguistically concise language (Grove et al., 2013; Waltz et al., 2010).

Step III: Scaling of Items

Once an item pool is generated, researchers should select the type of scale for developing the response set. In healthcare and nursing research, several types of scales are used which are broadly categorized as categorical and continuous scales. This categorization is based on the type and nature of a construct (Streiner et al., 2015). For example, participants' knowledge about diabetes mellitus can be measured on a categorical scale, whereas compassion and empathy should be measured on a continuous scale. A categorical scale includes a response set: "yes" or "no". However, continuous scales are developed using any of the three approaches: direct, comparative, and econometric estimation.

In direct estimation, the participants indicate their response by a check on an item or in a box. This approach includes visual analog, adjectival and Likert scales (Streiner et al., 2015). In visual analog scales, researchers develop a linear response set with two or three response options (Waltz et al., 2010). For example, a pain scale with a response set of "mild", "moderate", or "severe". The adjectival and Likert type scales are 4 to 7-point scales. However, the only difference between these two scales is that the former scale is unipolar (Streiner et al., 2015). For example, a researcher measures the overall satisfaction rate of patients from nursing care on a 4-point adjectival scale. The response set could be: "excellent", "very good", "good", and "fair". However, this overall rating

on a 5-point Likert scale could be: “excellent”, “good”, “undecided”, “poor”, and “dreadful”. This last example reflects the unipolar and bipolar nature of adjectival and Likert scales.

In the comparative scale development approach, the scale is designed in such a way that participants are given an opportunity to compare one object with another (Sridhar, 2016; Streiner et al., 2015). For example, if a researcher wants to construct a comparative scale to measure coping strategies of nurses, he or she can provide a list of coping strategies and ask nurses to rank these strategies based on their priorities in using those strategies.

During econometric scale development approach, a researcher designs a scale in such a way that the participants are expected to rate an item on two extremes (Streiner et al., 2015). For example, if a researcher wants to determine the effect of some nursing interventions on the independence level of patients, the scale could include a list of nursing interventions with two extremes (independence to dependence).

Step IV: Item Selection

The next step after scaling of generated items is item selection which is essential for operationalizing the intended construct adequately. Streiner et al., (2015) suggested two main criteria for selecting items; that is, interpretability and homogeneity. Interpretability means examining items in terms of ambiguity, double barreling, reading level of target population, jargons, and value laden wording. Ambiguity of items is determined by assessing the wording and vagueness of items. Double barrel items ask two questions at a time and therefore such items should be assessed by examining the

number of questions asked in each item. Value laden wording means the excessive use of adverbs of frequency such as “often”, “seldom”, “sometimes” and so forth, which may be confusing to participants. The reading level of the items of a scale should not be beyond the reading level of a 12-year-old except when the characteristics and educational level of the target population is known. Generally, interpretability of items is determined through the personal judgements of researchers and expert consultation (Streiner et al., 2015).

Homogeneity refers to the homogeneous nature of the items, that is, the items should be measuring different features of the intended construct and not different aspects of different constructs. It is determined through item-total correlation testing in which a researcher statistically determines the correlation of each item with the total scale by omitting the specific item. Prior to conducting item-total correlation analysis, the data is collected from a selected sample of the target population through a pilot study. The results are then analyzed using Pearson’s correlation coefficient. This testing is usually done as a part of psychometric testing. Omission of the specific item is necessary in order to avoid errors that could arise because of the omitted item’s correlation with itself. The acceptable item-total correlation of each item with the total scale should be more than 0.20 (Streiner et al., 2015).

Step V: Psychometric Testing

The psychometric testing of scale is the most essential step in the scale development process because it enables researchers to establish the reliability and validity of the constructed scale (Bolarinwa, 2015). By this time an adequate number of

items consistent with the dimensions of the construct has been determined, hence psychometric testing helps in further refining of the selected items.

Psychometric testing requires pilot studies and then large-scale validity and reliability studies. Pilot studies are designed to determine the feasibility of a future study and to perform preliminary psychometric testing of the developed scale (Thabane et al., 2010). However, large scale validity and reliability studies are often conducted after pilot studies. Generally, one pilot study is conducted (Streiner et al., 2015). The major difference between these two types of studies is the sample size. In pilot studies, the sample size ranges from at least 15 to 30 participants representative of the target population (Grove et al., 2013). However, for large scale studies the sample size should range from 100 to 250 subjects (Anthoine, Moret, Regnault, Sébille, & Hardouin, 2014). It has been suggested that for large scale reliability and validity studies, the number of participants should be selected in accordance with the number of items in the scale. The recommendations are that at least 10 to 20 participants per item should be selected (Anthoine et al., 2014; Grove et al., 2013). For example, if a scale contains 30 items the number of participants for large scale study should be 300.

Reliability refers to the stability and homogeneity of a scale in measuring the construct of interest (Kimberlin & Winterstein, 2008) where stability is the extent to which a scale measures consistent results on repeated occasions and homogeneity refers to correlation among the items and the construct of a scale. Simply put, homogeneity reflects the consistency of the conceptual and the operational definition (items of the scale) of the construct (Bolarinwa, 2015; Bannigan & Watson, 2009). Validity refers to

the extent to which a scale truly measures now the established operational definition of the construct of interest (Bannigan & Watson, 2009). There are several types of reliability and validity and different procedures to determine each type of reliability and validity. The most common types and methods of reliability are: *test-retest, inter-rater, parallel form, and internal consistency* (Streiner et al., 2015), and are discussed below.

Test-retest reliability. This type of reliability determines the ability of a scale to measure the construct of interest consistently over time (Kumar, 2016; Bannigan & Watson, 2009). In order to measure test-retest reliability the constructed scale is administered to the same participants at two different points in time (e.g., "T1" and "T2") and the responses are compared and correlated (Bolarinwa, 2015; Nimon, Zientek, & Henson, 2012). The interval between T1 and T2 should be not be too narrow nor too wide so as to avoid variations of scores and changes or errors in participants' responses (Bolarinwa, 2015; Kimberlin & Winterstein, 2008). Researchers' views about the minimum adequate interval vary, but a retest interval of 2 to 14 days is generally suggested (Streiner et al., 2015).

Inter-rater reliability. This type of reliability measures the extent to which different researchers collect consistent information using a similar scale under similar conditions (Bialocerkowski, 2008). The quantitative measure of inter-rater reliability or agreement between researchers is commonly calculated as Kappa coefficient which also "takes into account the agreement that could be expected by chance alone" (Kimberlin & Winterstein, 2008, p. 2288). The Kappa value of 0.60 or more is considered acceptable (Streiner et al., 2015).

Parallel form reliability. This type of reliability is similar to test-retest reliability in terms of method of scale administration. Unlike test-retest reliability, in the parallel form researchers use two different versions of the scale. However, it is important to ensure that the two different versions should not influence participants' responses (DeCoster, 2005). The obtained results are compared and correlated. The correlation coefficient value of more than 0.7 indicates good parallel form reliability (S. Bajpai & R. Bajpai, 2014).

Internal consistency. The internal consistency measures the extent of homogeneity among the items and the sub-scales of a scale (S. Bajpai & R. Bajpai, 2014). It is the most commonly used type of reliability (DeVellis, 2016). In order to measure this type of reliability, the scale scores are collected only one time and then analysed using different statistical tests such as inter-item correlation, Kuder-Richardson index, Cronbach's alpha, and split-half method (Bolarinwa, 2015). Among these methods, Cronbach's alpha is the most popular and commonly used test. The acceptable value of Cronbach's alpha ranges from 0.70 to 0.95 (Streiner et al., 2015; Bajpai & Bajpai, 2014; Tavakol & Dennick, 2011).

Validity

Again, as defined above, validity is the extent to which a scale truly measures now the established operational definition of the construct of interest. *Face, content, criterion, and construct validity* (Streiner et al., 2015; Bolarinwa, 2015) are most commonly considered when assessing scale validity.

Face validity. It is considered to be the most casual type of validity and includes an expert opinion and review of a scale and its items. Based on such a review a scale, a decision is made whether the scale is suitable for measuring the intended construct. The main limitation of face validity is that it is highly subjective (Bolarinwa, 2015).

Content validity. Similar to face validity, more than one expert or researcher evaluates a scale, its items, and content. The content refers to the conceptual definition of the construct, its attributes, and the selected items to operationalize the construct which is subject to expert evaluation. However, unlike face validity, content validity is more organized process. It is more structured and enables researchers to determine the level of consensus among experts that the scale is measuring the intended construct. This level of agreement is commonly presented in terms of Item-rated Content Validity Index (I-CVI) and Scale-level Content Validity Index (S-CVI). The acceptable value of I-CVI and S-CVI is more than 0.78 (Bolarinwa, 2015).

Criterion validity. Criterion validity refers to comparison and correlation of a scale with other available scales particularly with the gold standard instrument used to measure the intended construct. The comparison refers to the evaluation of the conceptual definitions of the construct in the selected scales while correlation refers to the comparison of item-total or inter-item correlation scores of the scales. It includes two types of validity namely, *concurrent* and *predictive* validity (Streiner et al., 2015). When assessing for concurrent validity, the researcher administers a scale and the gold standard instrument at the same time and the item scores are then compared and correlated. However, in predictive validity a scale is evaluated for its ability to predict any

prospective event, outcome, or result. Predictive validity gives an idea about the expected changes in scale scores of participants after certain time (Bolarinwa, 2015; Waltz et al., 2010). For example, a researcher measures the compassion levels of nursing students which may change over time due to students' experience, level of maturity, and gained knowledge. A scale which possess good predictive validity will indicate the changes in compassion levels when administered to the same group of students at the second time.

Construct validity. Construct validity refers to the ability of a scale to truly measure the intended construct using a hypothetical comparison with the construct (Streiner et al., 2015). This type of validity is measured on an ongoing basis by compiling research in which a scale has been used to measure the intended construct and the theoretical underpinnings on which the scale was constructed (Kimberlin & Winterstein, 2008). For example, if a researcher intends to measure the construct validity of a scale, developed on Watson's caring theory, he or she would develop different hypotheses, pertinent to the scale, from the theory. Those hypotheses will then be tested with the available literature on the usage of the scale. If the literature confirms the hypotheses then the scale is considered constructively valid. Therefore, based on the confirmed hypothesis testing, construct validity measures the degree of correlation between a scale and a theory that was used to develop the scale (Streiner et al., 2015). Another approach is to compile the findings of the studies in order to compare the psychometric findings of the scale over time and across regions. Construct validity further comprises discriminant and convergent validity. In discriminant and convergent validity, a researcher compares a scale with another instrument which should be constructed based on the same theoretical

underpinnings and then determines whether both the instruments are related and correlated. If the instruments are found to be related, then they are said to be convergently valid. If two instruments are not related they are said to be divergently valid (Streiner et al., 2015; Bolarinwa, 2015).

Step VI: Finalization of Scale

The last and final step of scale construction is the finalization of scale which is based on the validity and reliability testing (Waltz et al., 2010). The acceptable findings of the validity and reliability testing such as Cronbach's alpha value, expert consensus, item-total correlation and so forth, provides an indication that a scale measures the construct of interest. Therefore, researchers should carefully examine the results of psychometric testing. Some scales need further revision based on the results of pilot and validation studies and should be revised through the same process of scale construction (DeCoster, 2005).

Challenges in Constructing a Valid and Reliable Scale

During scale construction researchers could encounter many methodological challenges compromising the scale's reliability and validity (Miller, Reynolds, Ittenbach, Luce, Beauchamp, & Nelson, 2009). The first challenge that affects the validity and reliability of a scale is how to adequately conceptualize the construct of interest. Often, inappropriate theory selection and inadequate and narrow literature review could result in poor conceptualization (DeVellis, 2016) and poor operationalization of the construct with no predictive power (Miller et al., 2009). The second challenge is how to clearly differentiate the construct of interest from other similar constructs because poor

differentiation directly affects the construct validity of the scale. The third challenge is how to conceptualize and operationalize the construct of interest in a broader context because a too narrowly conceptualized construct limits applicability and generalizability of the scale (Miller et al., 2009). These three challenges can be tackled by defining more accurately the construct of interest from multiple viewpoints. Simply put, use of more than one method of conceptualization and operationalization of a construct increases the likelihood of constructing a valid, reliable, applicable, and broad scale. Therefore, researchers should use multiple approaches to construct conceptualization and item selection (Streiner et al., 2015).

Conclusions

Scales are commonly used methods of data collection in survey research which is an important subset of nursing research. Constructing and validating a scale is a complicated and tedious process. Therefore, prior to scale construction researchers should conduct an in-depth literature review to identify any existing scales measuring the construct of interest. If any existing scales are found pertinent to one's purpose then those scales should be adopted in order to save time and energy. However, if it is determined that another scale is needed then researchers should follow a systematic and stepwise approach such as outlined in this paper. It is up to the researchers' judgement to select single or multiple approaches to conceptualize and operationalize a construct and to select the appropriate type of measures of reliability and validity. However, it is important to use multiple approaches within each step to develop a scale that is applicable

for the intended study population as well as valid and reliable in measuring the given construct of interest.

References

Anthoine, E., Moret, L., Regnault, A., Sébille, V., & Hardouin, J. B. (2014). Sample size used to validate a scale: A review of publications on newly-developed patient reported outcomes measures. *Health and Quality of Life Outcomes*, *12*(1). 1-10. doi: 10.1186/s12955-014-0176-2

Artino A. R., Jr., La Rochelle, J. S., Dezee, K. J., & Gehlbach, H. (2014). Developing questionnaires for educational research: AMEE Guide No. 87. *Medical Teacher*, *36*(6), 463-474.

Avella, J.R. (2016). Delphi panels: Research design, procedures, advantages, and challenges. *International Journal of Doctoral Studies, 11*, 305-321. Retrieved from http://www.informingscience.org/Publications/3561

Bajpai, S., & Bajpai, R. (2014). Goodness of measurement: Reliability and validity. *International Journal of Medical Science and Public Health, 3* (2), 112-115. doi: 10.5455/ijmsph.2013.191120133

Balasubramanian, N. (2012). Likert technique of attitude scale construction in nursing research. *Asian Journal of Nursing Education and Research*, *2*(2), 65-69.

Bannigan, K., & Watson, R. (2009). Reliability and validity in a nutshell. *Journal of Clinical Nursing*, *18*(23), 3237-3243.

Bialocerkowski, A. E. (2008). Measurement error and reliability testing: Application to rehabilitation. *International Journal of Therapy & Rehabilitation*, *15*(10), 422-427.

Bolarinwa, O. A. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journal*, *22*(4), 195.

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health*, *27*(3), 281-291.

DeCoster, J. (2005). Scale construction notes. [Online]. Retrieved from http://www.stat-help.com/scale_20100316.pdf

DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). Newbury Park, CA. Sage

Furr, M. (2011). *Scale construction and psychometrics for social and personality psychology*. Thousand Oaks, CA: Sage.

Giesen, D., Meertens, V., Vis-Visschers, R., & Beukenhorst, D. (2012). *Questionnaire development.* Heerlen, Netherlands: The Hague Publishers Retrieved from https://www.cbs.nl/nr/rdonlyres/99b7482f-e09e-4d86-90a1-666dc42745bd/0/2012questionairedevelopmentart.pdf

Grove, S. K., Burns, N., & Gray, J. (2013). *The practice of nursing research: Appraisal, synthesis, and generation of evidence* (7th ed.). St. Louis, MO. Elsevier.

Hartrick-Doane, G., & Varcoe, C. (2015). *How to nurse: Relational Inquiry with individuals and families in changing health and health care contexts*. Philadelphia, PA: Lippincott.

Hessel, J. A. (2009). Presence in nursing practice: A concept analysis. *Holistic Nursing Practice*, *23*(5), 276-281.

Holosko, M. J., Thyer, B. A. (2011). *Pocket glossary for commonly used research terms*. Thousand Oaks, CA. Sage

Houser, J. (2016). *Nursing research: Reading, using and creating evidence*. (4th ed.). Burlington, MA. Jones & Bartlett Learning.

Kelley, K., Clark, B., Brown, V., & Sitzia, J. (2003). Good practice in the conduct and reporting of survey research. *International Journal for Quality in Health Care*, *15*(3), 261-266.

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal Health System Pharmacy*, *65*(23), 2276-2284.

Kumar, A. (2016). Review of the steps for development of quantitative research tools. *Advanced Practices in Nursing*, *1*(1), 1-7. doi: 103. doi:10.4172/apn.1000103

Leininger, M. (1993). Culture care theory: The comparative global theory to advance human care nursing knowledge and practice. In D. Gaut (Ed.), *A global agenda for caring* (pp. 3-19). New York, NY: National League for Nursing Press.

LoBiondo-Wood, G., Haber, J. (2014). *Nursing Research-E-Book: methods and critical appraisal for evidence-based practice* (8th ed.). St. Louis MO. Elsevier Health Sciences.

Masadeh, M. (2012). Focus group: reviews and practices. *The Journal of Applied Science and Technology*, *2*(10), 63-68.

Miller, V. A., Reynolds, W. W., Ittenbach, R. F., Luce, M. F., Beauchamp, T. L., & Nelson, R. M. (2009). Challenges in measuring a new construct: Perception of voluntariness for research and treatment decision making. *Journal of Empirical Research on Human Research Ethics*, *4*(3), 21-31.

Nassar-McMillan, S.C., Borders, L.D. (2002). Use of focus groups in survey item development. *The Qualitative Report*, 7(1). Retrieved October 7, 2005, for http://www.nova.edu/ssss/QR/QR7- 1/nassar.html.

Nimon, K., Zientek, L. R., & Henson, R. K. (2012). The assumption of a reliable instrument and other pitfalls to avoid when considering the reliability of data. *Frontiers in Psychology*, *3*,102. doi: 10.3389/fpsyg.2012.00102

Polit, D. F., & Beck, C. T. (2014). *Essentials of nursing research: Appraising evidence for nursing practice* (4th ed.). Philadelphia, PA. Lippincott.

Radhakrishna, R. B. (2007). Tips for developing and testing questionnaires/instruments. *Journal of Extension*. *45*(1). Retrieved from https://www.joe.org/joe/2007february/tt2.php

Rattray, J., & Jones, M. C. (2007). Essential elements of questionnaire design and development. *Journal of Clinical Nursing*, *16*(2), 234-243.

Sridhar, M. S. (2016). Research methodology: Measurement and scaling techniques [PowerPoint Slides] Retrieved from Online Web site: http://rizkia.staff.telkomuniversity.ac.id/files/2016/08/RM-6.pdf

Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use*. Oxford, UK: Oxford University Press.

Swanson, K. M. (1991). Empirical development of a middle-range theory of caring. *Nursing Research*, *40*, 161-166.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53-55. doi: 10.5116/ijme.4dfb.8dfd

Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., ... & Goldsmith, C. H. (2010). A tutorial on pilot studies: The what, why and how. *BMC Medical Research Methodology*, *10*(1), 1-10. doi: 10.1186/1471-2288-10-1

Timmins, F. (2015). Surveys and questionnaires in nursing research. *Nursing Standard*, *29*(42), 42-50.

Tools of research. (n.d.). Retrieved from http://results.mu.ac.in/myweb_test/Research%20Methadology-Paper-3/Chapter-9.pdf

Waltz, C. F., Strickland, O. L., & Lenz, E. R. (Eds.). (2010). *Measurement in nursing and health research*. New York, NY: Springer.

Watson, J. (1979). *The philosophy and science of caring*. Boston, MA: Little Brown.

Appendix II

Integrative Literature Review

An Integrative Literature Review to Develop the Relational Inquiry Capacities Scale

Hartrick-Doane and Varcoe (2015) proposed their Relational Inquiry approach can serve as a guide for practice for both nurses and nursing students. This approach was developed within the Canadian context and entails two components: "relational consciousness" and "inquiry as a form of action" (p. 3). Relational consciousness refers to the expectation that the nurse's role entails deliberately examining relational experiences of patients, families and other people involved in any given situation in order to understand the factors influencing the clinical situations. Inquiry as a form of action refers to the significance of critically examining each nursing situation and understanding all the influential factors and perspectives of the persons involved (Hartrick-Doane & Varcoe, 2015).

In order to practice from the Relational Inquiry perspective, nurses need to expand their thinking; that is, they need to reflect on patient situations and the context which the co-developers have categorized as layers of interpersonal, intrapersonal, and contextual factors which shape nurse-patient interactions (Hartrick-Doane, 2014; Hartrick-Doane & Varcoe, 2015). The co-developers explained that the ability to intervene and provide effective nursing care depends on the relational capacities: "compassion", "self-compassion" "curiosity", "commitment", "competence", and "correspondence" (Hartrick-Doane & Varcoe, 2015). The Relational Inquiry approach has been purported as an effective guide for nursing educators for bridging the gap between nursing theory and practice by helping students integrate theoretical nursing knowledge in practice. This approach has also been claimed to be critical for improving

students' clinical learning and can serve as a means to assess the capacity of students to intervene. That is, nursing educators require ways to assess student levels of the six relational capacities. The purpose of this literature review was twofold: 1) to select relevant questionnaires or scales for measuring the capacities: compassion, curiosity, commitment, competence, and correspondence, 2) to analyze these scales through the lens of the Relational Inquiry approach, 3) to compile the definitions of the capacities. The scales and definitions will then be the basis for developing a new instrument: the RICS. Before proceeding, I provide theoretical and empirical support for the effectiveness of Relational Inquiry based teaching.

Relational Inquiry Based Teaching

Relational Inquiry based teaching could be described as a process for assisting nursing students to translate theoretical nursing knowledge, skills, values, and ideologies into effective nursing actions (Hartrick-Doane & Varcoe, 2015). The effectiveness of the Relational Inquiry approach has not been studied extensively in nursing education research. However, several authors have considered this approach as a guide for nursing educators who want to revisit their teaching and learning pedagogies. For example, Hartrick-Doane (2002a) argued that behavioral teaching pedagogies such as traditional lectures, skill performance, and so forth, can only assist nursing students to learn technical skills because such pedagogies only cover the course content. Behavioral teaching pedagogies do little to assist students to develop capacities for building meaningful relationships among themselves and with their patients. Therefore, it is recommended that nursing educators should use Relational Inquiry as an education tool

given there is empirical evidence that Relational Inquiry enable nursing students to gain an in-depth understanding of nursing concepts (Hartrick-Doane, 2002a). Consistent with this notion, Hartrick-Doane (2002b) indicated that behavioral teaching pedagogies cannot prepare creative and caring nurses because such pedagogies only help educators to develop technological and empirical skills of nurses. Hartrick-Doane moved beyond technical skills when teaching nursing ethics, for example, by engaging students in activities which draw their attention to morally complex and uncertain nursing situations. She was able to foster student creativity and caring abilities and prepare students for ethical nursing practice. Similarly, Spadoni, Doane, Sevean, and Poole (2015) adapted the Relational Inquiry approach to develop a 6-week educational project comprised of mask making and storytelling. The purpose of their project was to help students understand the essence of caring in nursing practice. Through participation in this project, the students recognized the importance of caring and connecting with self and others. They were able to improve their self-reflection and awareness skills. The realized that in order to become a caring individual, they should be honest about their thoughts about themselves and each other, practice active listening, and analyze their thoughts and feelings from various different perspectives.

This study and the claims made by Hartrick-Doane highlights the importance of Relational Inquiry approach as a guide to develop effective teaching and learning pedagogies and Relational Inquiry can serve as an important educational tool. However, these studies provide little evidence of the use of Relational Inquiry approach in nursing practice, moreover, whether nursing students are able to adopt this approach for their

clinical practice. Given that the Relational Inquiry approach is useful for guiding nursing practice and advancing student learning, it is important to explore the Relational Inquiry nursing capacities: compassion, competence, commitment, correspondence, and curiosity in educational based research.

Significance of the Relational Nursing Capacities

The relational capacities, with the addition of *self-compassion*, are essential for enabling nursing students to critically assess any given clinical situation; examine factors influencing the situation; and then discern their nursing obligations and responsive actions (Hartrick-Doane & Varcoe, 2015). Researchers across the world have highlighted importance of these capacities if one considers the wide range of literature about the capacities: compassion, commitment, and competence (Burnell, & Agan, 2013; Lee & Seomun, 2016a, 2016b; Papadopoulos & Ali, 2016; Yang & Jiang, 2014). However, missing in the literature is research substantiating to what degree nurses possess these capacities and whether these capacities have been operationalized as an interrelated whole. This dearth of literature may exist due to limited instruments measuring any of these six capacities (Burnell, & Agan, 2013; Lee & Seomun, 2016b; Papadopoulos, Shea, Taylor, Pezzella, & Foley, 2016; Yang & Jiang, 2014).

The need for instruments is consistent with the findings drawn from an integrative review by Papadopoulos and Ali (2016) who focused on competence and compassion. The authors recommended the need to develop instruments for measuring compassion and other interrelated concepts. Although this integrative review only highlighted research concerning compassion and competence, it can be implied that the other

interrelated concepts: curiosity, commitment, and correspondence should also be measured. Therefore, this integrative literature review was conducted to develop an item pool for Relational Inquiry Capacities Scale (RICS).

Literature Search

A comprehensive literature search was performed within the databases: CINHAL, Google Scholar, PubMed, and Science Direct using specific keywords and phrases. The specific key words were: “compassion questionnaire/scale”; “curiosity questionnaire/scale”; “nursing commitment questionnaire/scale”; “competence questionnaire/scale”; and “nursing and professional correspondence questionnaire/scale”. The key phrases were: “measuring compassion of nurses and nursing students”; “levels of compassion of nursing students and nurses”; “measuring commitment of nurses and nursing students”; and “levels of commitment of nursing students and nurses”; “measuring competence of nurses and nursing students”; “levels of competence of nursing students and nurses”; “measuring curiosity of nurses and nursing students”; “levels of curiosity of nursing students and nurses”; “measuring correspondence of nurses and nursing students”; and “levels of correspondence of nursing students and nurses”.

Literature Search, Inclusion and Exclusion Criteria, and Study Selection

The initial search retrieved 14,190 articles (CINHAL (n= 936), PubMed (n=57), Google Scholar (n= 11, 890), and Science Direct (n=1307)). The search was then limited to journal articles, questionnaire and scale development and validation studies and repetitive results were screened resulting in 945 articles. Then, 865 articles were excluded after screening for relevant titles. The 865 excluded articles were focused on

psychometric testing of a variety of scales (e.g., empathy, love compassion, and organizational commitment and so forth) and different models and theories related to the six capacities, and some were descriptive studies and literature reviews. The final selection of the articles for this review was based on the following criteria: original research studies published in English focusing on questionnaire development or using a questionnaire to measure any of the six relational capacities in nursing students or professional nurses. The literature reviews, position statements, letters, discussions, and editorials were excluded. The conceptual papers were excluded because the intention was to compare both conceptual and operational definitions of the relational capacities and operational definitions would not have been available in conceptual papers. An additional search of the reference lists of the selected articles was performed to identify further articles but this resulted no additional articles. The remaining 80 articles which met the inclusion criteria were screened again after reading the abstracts. This final screening excluded another 46 articles. Out of the remaining 34 articles, one article's full text that outlined the method for developing the professional commitment scale could not be retrieved. The final 33 full-text articles were read and only 15 articles were selected because they fulfilled the inclusion criteria.

Overview of the Included and Reviewed Studies

The 15 studies selected were quantitative in nature and used either cross-sectional or survey research design. These studies were conducted in countries such as Finland, Jordan, China, Japan, and the US. The main purpose of the studies was to develop a questionnaire or scale for measuring the relational capacities or to determine the levels of

these capacities in nursing students and professional nurses. In total, six articles focused on the development or psychometric testing of scales to measure compassion (Lee & Seomun, 2016; Lown, Muncer, & Chadwick, 2015; Burnell & Agan, 2013; Raes, Pommier, Neff, & Van Gucht, 2011; Pommier, 2010; Neff, 2003), eight articles outlined the process for development of nursing competence scales (Nilsson et al., 2014; Hsu & Hsieh, 2013; Takase & Teraoka, 2011; Safadi, Jaradeh, Bandak,& Froelicher, 2010; Lin, Hsu, Li, Mathers, & Huang, 2010; Cowan, Wilson-Barnett, Norman, & Murrells, 2008; Liu, Kunaiktikul, Senaratana, Tonmukayakul, & Eriksen, 2007; Meretoja, Isoaho, & Leino-Kilpi, 2004). One article was focused on measuring nurses' commitment (Lin, Wang, Li, & Huang, 2007) however the full text article and scale used in this research could not be found. One article was focused on the development of a scale to measure curiosity (Kashdan et al., 2009). There were no articles on scale development for measuring nurses' correspondence.

Critical Appraisal of the Included Studies

Consistent with the PHAC (2014) critical appraisal toolkit, the researchers of the included studies met most of the criteria and the overall strength of the studies was moderate to strong. In each article, the research question, purpose, target population, sample and its characteristics were clearly outlined. Most of the researchers were guided by a theoretical and conceptual framework. However, conceptualization of the measured constructs was not evident in a few studies (Burnell & Agan, 2013; Lown et al., 2015; Safadi et al., 2010; Cowan et al., 2008; Takase & Teraoka, 2011; Nilsson et al., 2014; Hsu & Hsieh, 2013). The sample size for the validity and reliability testing of the

questionnaires and scales ranged from 250 to 1534 and the samples were representative of the target population. For pilot studies, at least 10% sample of the actual sample was used. In all the studies, ethical approval and informed consent was obtained and essential measures were taken to ensure the confidentiality and privacy of the participants. All of the studies used robust and systematic process for questionnaire and scale development and appropriate statistical procedures were employed to test for validity and reliability. All the questionnaires and scales were valid and reliable except the short form self-compassion scale developed by Raes et al., (2011). This short form self-compassion scale has low internal consistency value (0.55). A major limitation of the included studies was that the researchers studied the relational capacities as single entities, rather than as a whole. According to Hartrick-Doane and Varcoe (2015) these six capacities complement each other and should be applied as an interdependent unit. The detailed account of the critical appraisal results of the 15 studies can be found in Appendix I.

Comparison of the Definitions of the Six Constructs

Through this literature review, I was able to make comparisons between the extant definitions of the six capacities: compassion, self-compassion, curiosity, commitment, competence, and correspondence with how they are conceptualized in the Relational Inquiry approach (Hartrick-Doane & Varcoe, 2015). The Relational Inquiry comparisons are discussed below.

Compassion and Self-Compassion

Compassion was defined in similar ways in the literature selected. Most definitions were somewhat consistent, that is, missed at least one or more aspects with

Relational Inquiry (Lee & Seomun, 2016; Lown et al., 2015; Raes et al., 2011; Pommier, 2010; Neff, 2003). Burnell and Agan's (2013), however, defined compassion differently.

It is interesting to note that Neff (2003) also conceptualized self-compassion from a Buddhist perspective and included three components namely, *self-kindness, humanity, and mindfulness*. The author defined self-compassion as "being open to and moved by one's suffering, experiencing feelings of caring and kindness towards oneself, taking an understanding non-judgmental attitude towards one's inadequacies and failures, and recognizing that one's experience is a part of the common human experience (p. 224). Hartrick-Doane and Varcoe's (2015) also discuss self-compassion as "consciously and intentionally choosing to act and respond to yourself, others, and the situation to promote well-being at interpersonal, intrapersonal, and contextual levels" (p.113). The comparison of two definitions reveals two similarities. First, both Neff and Hartrick-Doane and Varcoe conceptualized self-compassion as a caring attitude towards one self and one's recognition of personal limitations. Second, the authors grounded self-compassion within the context of mindfulness and deliberate examination of personal feelings and thoughts. The main difference between these two definitions is the emphasis on self-refection and critical reflection for the purposes of examining a given nursing situation. Hartrick-Doane and Varcoe's (2015) definition focused on self-reflection of one's actions and deliberately revisiting one's actions. However, Neff's definition did not explicitly highlight the need for self-reflection of one's actions.

Pommier (2010) adopted the definition of compassion by Neff (2003): "being open to and moved by the suffering of others so that one desires to ease their suffering. It

also involves offering patience, kindness, non-judgmental understanding, and recognizing that all humans are imperfect and make mistakes" (Neff, 2003, p. 224). This definition was somewhat consistent with the Relational Inquiry approach because, unlike Hartrick-Doane and Varcoe, Pommier emphasized easing one's suffering. Hartrick-Doane and Varcoe (2015) did not emphasize taking actions to ease one's suffering rather they stressed that one should "share suffering by being in solidarity with persons and doing with one another" (pp. 103-104). These two definitions were similar because of the inclusion of traits such as kindness, understanding, and recognizing one's limitations as indicative of compassion. Raes's et al., (2011) also aligned with Neff's definition and their scale item were actually extracted from the work of Pommier (2010).

Lown et al., (2015) did not provide a clear definition of compassion. Also, the authors' scale intended to measure compassionate healthcare practice rather than compassion levels of healthcare professionals. Therefore, only Lown's et al., 2015) compassionate healthcare scale was analyzed through the lens of the Relational Inquiry approach. Lee and Seomun (2016) developed a scale to measure nurses' compassion competence. The term compassion competence was conceptualized as: "communication (expressing understanding and compassion towards patients and families), sensitivity (ability to recognize), and insight (ability to clearly understand patients and their needs" (p. 80). This conceptualization was somewhat relevant because compassion and competence were considered to follow an interdependent relationship.

As previously stated, Burnell and Agan's (2013) definition was inconsistent with the Relational Inquiry approach. The authors developed an assessment tool to measure

compassionate care attributes. The authors did not provide a conceptual definition of compassion and the assessment tool was mainly centered on spiritual nursing care.

Competence

Out of eight research studies that focused on scale development to measure competence, six researchers' definitions (Meretoja et al., 2004; Liu et al., 2007; Cowan et al., 2008; Safadi et al., 2010; Takase &Teraoka, 2011) were found to be somewhat consistent with the Relational Inquiry approach while three researchers' (Nilsson et al., 2014; Hsu & Hsieh, 2013; Lin et al., 2010) definitions were inconsistent.

Meretoja et al., (2004) developed the nurses' competence scale based on Benner's (1984) Novice to Expert model and its underlying theoretical assumptions. The authors did not provide an explicit definition of competence, but operationalized competence into seven categories: helping, teaching, diagnostic functions, managing situations, therapeutic interventions, ensuring quality, and work role that, together, reflected practical application which is espoused in the Relational approach.

Liu et al., (2007) developed the competence inventory for nurses in China. The term competence was conceptualized and operationalized after conducting a preliminary study. Based on the findings, eight dimensions of competence were developed which included leadership, clinical care, interpersonal, ethical care, teaching, professional development, critical thinking, and research aptitude. The authors did not provide the conceptual definition of competence. Cowan et al., (2008) developed a self-assessment tool for measuring nursing competence. The definition of competence was operationalized in terms of nurses' knowledge, skills, professional judgement, ethics,

values, reflective and context dependent practice. Though the conceptual definition was not consistent with the Relational Inquiry approach, some scale items would be suitable for inclusion in an item pool for RICS. Safadi et al., (2010) developed a scale to measure competence levels of nursing students. The authors did not provide a conceptual definition of competence but operationalized it in terms of five competencies: management, professional development, nursing process, problem solving and knowledge. The operational definition was somewhat consistent with the Relational Inquiry approach because of its focus on problem solving and knowledge. Takase and Teraoka (2011) developed the holistic nursing competence scale. After conducting a concept analysis, conceptualized competence as: nurses' knowledge, values, professionalism, and motivation to provide nursing care which were somewhat consistent with the Relational Inquiry approach.

Hsu and Hsieh (2013) developed the competence inventory to measure nursing students' learning outcomes. Competence was conceptualized and operationalized based on the eight core competencies proposed by the American Nurses Association. These competencies were: ethical practice, clinical skills, continuous learning, biomedical knowledge, caring and critical thinking. This definition was consistent with the Relational Inquiry approach because it focused on nursing ethical values, caring and critical thinking skills, and student learning. Nilsson et al., (2014) developed nurses' professional competence scale. The authors operationalized competence in terms of nursing care, values based care, teaching, education, legal practice, medical care, and documentation. This definition was inconsistent with the Relational Inquiry approach

because it was focused on psychomotor skills and biomedical knowledge of students. Though the Relational Inquiry approach highlights the importance of biomedical knowledge and psychomotor skills, it places greater emphasis on the relational skills of nurses and students. Nilsson et al., (2014) did not operationalize relational knowledge and skill development. Lin et al., (2010) developed public health nurses' professional competence scale. Competence was operationalized as: basic care, community based, self-competence, and teaching competence. However, the conceptual and operational definitions were inconsistent with the Relational Inquiry approach.

Curiosity

Only one scale was found to measuring curiosity levels of undergraduate students. Kashdan et al., (2009) developed the Curiosity and Exploration Inventory II (CEI-II). The authors' conceptualized curiosity in terms of three traits: "active information and opportunity seeking behavior, willingness to accept uncertainty in life, and tolerance of uncertainty" (p. 989). In contrast, Hartrick-Doane and Varcoe (2015) underscored that "being curious is about being interested, inquisitive, and open to uncertainty that is part of disease and illness experiences. It is the capacity to work between knowing and not knowing" (p. 115). After comparison, Hartrick-Doane and Varcoe (2015) emphasized that accepting uncertainty in one's life is an essential attribute of being curious because it provides oneself with the opportunity to improve learning. The ideology of CEI-II was based on determining students' motivation, learning, and well-being and missed the attribute of uncertainty. Also, Hartrick-Doane and Varcoe (2015) conceptualized

curiosity in terms of nurse patient relational relationship and willingness to learn more about patient's perspectives which was missing in Kashdan's et al., (2009) definition.

Commitment and Correspondence

No scales for measuring commitment and correspondence were found in my review of literature.

Description of the RICS and Item Pool

As aforementioned I conducted the literature review to develop an item pool for RICS. Five of Neff's (2003) Self-Compassion scale items and five of Pommier's (2010) compassion scale items were consistent with the Relational Inquiry approach. These items were extracted and adapted for the item pool of the RICS self-compassion subscale. Raes's et al., (2011) short form Self-Compassion scale was developed from Neff's (2003) Self-Compassion scale and Pommier's (2010) et al., (2011) scale. Therefore, none of the items from this scale were included in the RICS item pool. Three items of the Compassionate Healthcare scale (Lown et al., 2011) and on item in the Compassion Competence scale (Lee & Seomun, 2016) were consistent with the Relational Inquiry approach, but were not included in the item pool for RICS because the authors did not permit extraction or adaptation of the items. Burnell and Agan's (2013) Compassionate Care Assessment Tool items were inconsistent with the Relational Inquiry approach because related to the attribute of caring, patient's need for compassion and the nurses' ability to be compassionate. Moreover, the scale was developed from the Spiritual Needs Survey and Coping Behavior Inventory. These two instruments were inconsistent with

the Relational Inquiry approach because Hartrick-Doane and Varcoe do not explicitly address spiritual care and helping patients to cope as nursing domains.

Out of the competence scales, five scales contained items that were in line with the Relational Inquiry Approach (Meretoja et al., 2004; Liu et al., 2007; Hsu & Hsieh, 2013; Safadi et al., 2010; Takase & Teraoka, 2011), but items were extracted and adapted from only three scales because Liu et al., (2007) and Hsu and Hsieh (2013) did not permit extraction and adaptation of items. The critical examination of the Meretoja's et al., (2004) nurses' competence scale found that five items in three categories: helping, managing situations, and therapeutic interventions, were somewhat relevant with the Relational Inquiry approach. For the RICS curiosity subscale, some items of Kashdan's et al., (2009) CEI-II inventory were relevant to the Relational Inquiry approach and were included in the item pool. Since no scales for measuring commitment and correspondence could be found the item pool was developed from the assumptions of the Relational Inquiry approach.

In total, 21 items were extracted and adapted from the reviewed scales. After careful rewording of the items, the items were compiled for the RICS. The remaining items, not found during the literature review, were developed by drawing from Relational Inquiry. Finally, a preliminary draft of the RICS was developed which comprises six subscales and 71 items. The five subscales are: compassion (compassion for others (15 items) and self-compassion (13 items), curious (9 items), commitment (11 items), competence (16 items), and correspondence (9 items). The response set consists of 6-point Likert scale: "strongly agree", "agree", "slightly agree", "slightly disagree",

disagree, and “strongly disagree”. The complete item pool for the RICS is presented in Appendix II.

Conclusions

Based on the results of this integrative literature review, it is evident that limited scales exist measuring the only three (i.e., compassion, competence, and curiosity) of these relational capacities. Commitment and Correspondence scales located in the review were not pertinent to nursing students rather involved measuring career satisfaction and choices, professional commitment, professional development, and organizational commitment of nurses. Therefore, such scales were excluded.

The comparisons of the identified scales their extant definitions of the five capacities indicated that some of the scales were consistent and others were somewhat consistent with the Relational Inquiry approach. However, the items of the scales were inconsistent. Therefore, a preliminary draft of the RICS was developed which included selected items from the reviewed scales and remaining items were developed from the assumptions of Relational Inquiry approach.

References

Burnell, L., & Agan, D. L. (2013). Compassionate care: Can it be defined and measured? The development of the compassionate care assessment tool. *International Journal of Caring Sciences*, *6*(2), 180-187.

Cowan, D. T., Wilson-Barnett, D. J., Norman, I. J., & Murrells, T. (2008). Measuring nursing competence: Development of a self-assessment tool for general nurses across Europe. *International Journal of Nursing Studies*, *45*(6), 902-913.

Hartrick-Doane, G. (2014). Beyond relational oblivion. *Advances in Nursing Science*, *37*(2), 84-86.

Hsu, L. L., & Hsieh, S. I. (2013). Development and psychometric evaluation of the competency inventory for nursing students: A learning outcome perspective. *Nurse Education Today*, *33*(5), 492-497.

Kashdan, T. B., Gallagher, M. W., Silvia, P. J., Winterstein, B. P., Breen, W. E., Terhar, D., & Steger, M. F. (2009). The curiosity and exploration inventory-II: Development, factor structure, and psychometrics. *Journal of Research in Personality*, *43*(6), 987-998.

Lee, Y., & Seomun, G. (2016a). Compassion competence in nurses. *Advances in Nursing Science*, *39*(2), E54-E66.

Lee, Y., & Seomun, G. (2016b). Development and validation of an instrument to measure nurses' compassion competence. *Applied Nursing Research*, *30*, 76-82.

Lin, C. J., Hsu, C. H., Li, T. C., Mathers, N., & Huang, Y. C. (2010). Measuring professional competency of public health nurses: Development of a scale and psychometric evaluation. *Journal of Clinical Nursing*, *19*(21-22), 3161-3170.

Lin, C. J., Wang, H. C., Li, T. C., & Huang, L. C. (2007). Reliability and validity of nurses' job satisfaction scale and nurses' professional commitment. *Mid-Taiwan Journal of Medicine*, 12, 65-75.

Liu, M., Kunaiktikul, W., Senaratana, W., Tonmukayakul, O., & Eriksen, L. (2007). Development of competency inventory for registered nurses in the People's Republic of China: Scale development. *International Journal of Nursing Studies*, *44*(5), 805-813.

Lown, B. A., Muncer, S. J., & Chadwick, R. (2015). Can compassionate healthcare be measured? The Schwartz Center compassionate care scale. *Patient Education and Counseling*, *98*(8), 1005-1010. doi: 10.1016/j.pec.2015.03.019

Meretoja, R., Isoaho, H., & Leino-Kilpi, H. (2004). Nurse competence scale: Development and psychometric testing. *Journal of Advanced Nursing*, *47*(2), 124-133.

Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity*, *2*(3), 223-250.

Nilsson, J., Johansson, E., Egmar, A. C., Florin, J., Leksell, J., Lepp, M., ... & Carlsson, M. (2014). Development and validation of a new tool measuring nurses self-reported professional competence—The Nurse Professional Competence (NPC) Scale. *Nurse Education Today*, *34*(4), 574-580.

Papadopoulos, I., & Ali, S. (2016). Measuring compassion in nurses and other healthcare professionals: An integrative review. *Nurse Education in Practice*, *16*(1), 133-139.

Papadopoulos, I., Shea, S., Taylor, G., Pezzella, A., & Foley, L. (2016). Developing tools to promote culturally competent compassion, courage, and intercultural communication in healthcare. *Journal of Compassionate Health Care*, *3*(1), 1-10. doi: 10.1186/s40639-016-0019-6

Pommier, E. A. (2010). The compassion scale. Dissertation Abstracts International Section A: Humanities and Social Sciences, 72, 1174.

Raes, F., Pommier, E., Neff, K. D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the self-compassion scale. *Clinical Psychology & Psychotherapy*, *18*(3), 250-255.

Safadi, R., Jaradeh, M., Bandak, A., & Froelicher, E. (2010). Competence assessment of nursing graduates of Jordanian universities. *Nursing & Health Sciences*, *12*(2), 147-154.

Takase, M., & Teraoka, S. (2011). Development of the holistic nursing competence scale. *Nursing & Health Sciences*, *13*(4), 396-403.

Yang, G. F., & Jiang, X. Y. (2014). Self-directed learning readiness and nursing competency among undergraduate nursing students in Fujian province of China. *International Journal of Nursing Sciences*, *1*(3), 255-259.

Hartrick-Doane, G., & Varcoe, C. (2015). *How to nurse: Relational Inquiry with individuals and families in changing health and health care contexts*. Philadelphia, PA: Lippincott.

Spadoni, M., Doane, G. H., Sevean, P., & Poole, K. (2015). First-year nursing students—developing relational caring practice through inquiry. *Journal of Nursing Education*, *54*(5), 270-275.

Hartrick-Doane, G. (2002a). Beyond behavioral skills to human involved processes: Relational nursing practice and interpretive pedagogy. *Journal of Nursing Education*, *41*(9), 400-404.

Hartrick-Doane, G. (2002b). In the spirit of creativity: The learning and teaching of ethics in nursing. *Journal of Advanced Nursing*, *39*(6), 521-528.

Literature Summary Tables

Author/Year of Publication/ Purpose	Methods	Findings	Relevance/Irrelevance to the Relational Inquiry Approach	Strengths/Limitations
Studies on Compassion				
Neff (2003) Development and validation of a Self-Compassion Scale	Self-Compassion was conceptualized in Buddhist philosophy and included three components namely, self-kindness, humanity, and mindfulness. The study was completed in three phases with three samples: 391, 232, and 43. In the first phase, the participants' responses were used to develop the scale and in the second and third phase the scale's psychometric properties were assessed. The exploratory and confirmatory factor analyses were used for final item selection.	The final 20 item self-compassion scale was reliable (α=0.93) and valid (CFI=0.92).	Neff's definition of self-compassion was somewhat consistent with the Relational Inquiry approach because it did not capture the importance of self-reflection of one's actions in any given situation.	**Strengths** Systematic and robust method for scale development, appropriate statistical procedures, theoretically valid scale, and the use of large sample for validity and reliability testing. **Limitations** Self-compassion was measured as a separate entity from other similar constructs such as caring which could have influenced the construct validity.
Pommier (2011) Development and psychometric testing of the Compassion Scale	Compassion was conceptualized in the Buddhist concept of compassion and Neff's (2003) concept. It included three components namely, self-	The scale was valid (CFA of individual items >0.50) and reliable (Cronbach's alpha=0.90). Compassion	The definition of compassion was somewhat consistent with the Relational Inquiry approach because it conceptualized Neff's (2003) concept of	**Strengths** Systematic and robust method for scale development, appropriate statistical procedures, theoretically valid scale. **Limitations**

	kindness, humanity, and mindfulness. The scale development was completed in two phases with a sample of 439 and 510 participants. The confirmatory factor analysis, split half reliability, Cronbach's alpha, and convergent validity were used for reliability and validity testing.	was correlated with compassionate love, wisdom, social connectedness, and empathy providing support for convergent validity.	compassion and self-compassion. Moreover, the sample mainly consisted of white people which could have limited the generalization of items in other cultures which is inconsistent with the Relational Inquiry approach.	The sample mainly consisted of white people which could have limited the generalization of items to other ethnic groups.
Raes et al., (2011) Construction and factorial analysis of short form of Self-Compassion scale	Compassion was conceptualized in the Buddhist concept of compassion and Neff's (2003) concept. The scale was constructed in three phases with two samples (271 and 185) from Dutch population and one sample (415) from American population. Correlation analysis between self-compassion scale and short form scale was conducted.	The correlation analysis between self-compassion scale and short form scale resulted in r=0.97. The internal consistency of the scale's items ranged from 0.55 to 0.81. The CFA value was 0.97.	The definition of compassion was somewhat consistent with the Relational Inquiry approach because it was conceptualized on the Buddhist concept of compassion and Neff's (2003) concept and did not emphasize self-reflection on one's actions.	**Strengths** The scale was tested on two different populations through a systematic and robust process. **Limitations** The internal consistencies were low from 0.55 to 0.81 compared to original version.
Burnell & Agan (2013) Development of Compassionate Care Assessment tool	Compassion was conceptualized in terms of spiritual care, the attribute of caring, patient's need for compassion and the nurses' ability to be compassionate. The	The tool was valid (content validity) and reliable (Cronbach's alpha for the subscales ranged from 0.77 to 0.86).	The conceptualized and operationalized definition of compassion was inconsistent with the Relational Inquiry approach.	**Strengths** Conceptualization of compassion based on two previous surveys, validity and reliability measures were appropriate, confirmatory factor analysis was employed.

	final item development was based on two surveys: spiritual needs of patients and caring behavior of nurses. The tool was tested on 250 patients.			**Limitations** Lack of a theoretical framework for tool development and minimal validity testing.
Lown et al., (2015) Development and psychometric testing of the Compassionate Care Scale	The scale was pilot tested on a sample of 801 participants. The psychometric properties were assessed using Cronbach's alpha, exploratory and confirmatory analysis, and Mokken Analysis.	The scale was valid and reliable (α=0.97 and 0.95)	The conceptual definition of compassion was unclear. However, the items of the scale were somewhat relevant to the concept of compassion in the Relational Inquiry approach.	**Strengths** Systematic process for scale development and appropriate statistical analysis. **Limitations** Lack of a theoretical framework for tool development, lack of conceptual definition of compassion, and the scale was aligned towards compassionate behavior of doctors.
Lee and Seomun (2016) Development and psychometric testing of Nurses' Compassion Competence Scale	The term compassion competence was conceptualized as three main elements "communication (expressing understanding and compassion towards patients and families), sensitivity (ability to recognize), and insight (ability to clearly understand patients and their needs" (p.80). A rigorous concept analysis was used to develop the items of the scale. The	The scale was valid (item total correlation >0.30) reliable (α=0.91)	This definition of compassion was somewhat consistent with the Relational Inquiry approach because both definitions conceptualized compassion and nurses' competence in an interrelated relationship. Some items were consistent with the Relational Inquiry approach.	**Strengths** Systematic process for scale development and appropriate statistical analysis. **Limitations** The lack of cross cultural validation would limit the use of this scale.

	content, convergent and face validity was assessed. The reliability was assessed using Cronbach's alpha. The scale was pilot tested on a sample of 660 nurses.			
Studies on Competence				
Meretoja et al., (2004) Development and psychometric testing of Nurses' Competence Scale	The definition of competence was conceptualized from Benner's novice to expert model. It was operationalized into seven categories; helping, teaching, diagnostic functions, managing situations, therapeutic interventions, ensuring quality, and work role. The psychometric properties of the scale were assessed on a sample of 498 nurses. The measures such as content, concurrent, and construct validity and internal consistency were used for reliability and validity testing.	The scale was found to be valid (construct validity agreement of the expert > 50%) and reliable (the value of Cronbach's alpha was not given).	The conceptual definition was somewhat consistent with the Relational Inquiry approach. However, three categories (helping, managing situations, and therapeutic interventions) were found somewhat relevant to the conceptualization of competence in Relational Inquiry approach.	**Strengths** Systematic and rigorous process for scale development and the use of nursing model to conceptualize and operationalize the variables. **Limitations** The sample consisted of patients in the medical surgical units only.
Liu et al., (2007) Development and testing of Competence Inventory for Nurses.	The term "competence" was conceptualized and operationalized into 8 dimensions such as leadership, clinical care, interpersonal, ethical care,	The inventory was valid (construct, criterion (r=0.44), contrast group validity (p<0.001), and	The conceptual definition was inconsistent with the Relational Inquiry approach. However, several items from three categories (clinical care,	**Strengths** Systematic and rigorous process for scale development. **Limitations** The initial item pool was based on a qualitative study in the

	teaching, professional development, critical thinking, and research aptitude. The sample consisted of 815 nurses. The psychometric analysis was conducted using factor and item analysis, construct, criterion, and contrast group validity, and reliability.	reliability. and reliable (α=0.89).	interpersonal relations, and ethical practice) were found somewhat relevant to the conceptualization of competence in Relational Inquiry approach.	Chinese context therefore the scale may not be generalizable to other populations.
Cowan et al., (2008) Development of a self-assessment tool for measuring nursing competence of general nurses	The definition of competence was operationalized in terms of nurses' knowledge, skills, professional judgement, ethics, values, reflective and context-dependent practice. The sample consisted of 588 nurses. The psychometric analysis was conducted using factor and item analysis, construct and content validity, and reliability.	The scale was valid and reliable (α=0.961).	The conceptual definition was inconsistent with the Relational Inquiry approach. However, several items of the scale were found somewhat relevant to the conceptualization of competence in Relational Inquiry approach.	**Strengths** Systematic and rigorous process for scale development and the sample consisted of participants from five countries, namely UK, Spain, Greece, Germany, and Belgium. **Limitations** Lack of theoretical framework for conceptualization of competence.
Lin et al., (2010) Development and testing of public health nurses' Professional Competence Scale.	Competence was operationalized into four domains: basic care, community based, self-competence, and teaching	The scale was valid (content validity (indices >0.80) and reliable (α=0.93-0.97).	The conceptual definition was inconsistent with the Relational Inquiry approach and none of the items were pertinent to the	**Strengths** Extensive and systematic process for scale development and appropriate statistical analysis. **Limitations**

	competence. The sample consisted of 1534 public health nurses. The psychometric analysis was conducted using factor and item analysis, construct content, discriminant, and convergent validity. Reliability was assessed using Cronbach's alpha and inter-item correlation.		Relational Inquiry approach.	Lack of theoretical framework for conceptualization of competence and limited to a single population.
Safadi et al., (2010) Competence assessment of nursing students	Competence was operationalized in terms of five competencies: management, professional development, nursing process, problem solving and knowledge. The sample consisted of 258 nursing students. The psychometric properties were assessed through face validity and Cronbach's alpha.	The scale was found to be valid and reliable (α=0.97).	The conceptual definition was not stated. The operational definition was somewhat consistent with the Relational Inquiry and several items were pertinent to the Relational Inquiry approach.	**Strengths** First scale specifically focused on nursing graduates. **Limitations** Lack of theoretical framework for conceptualization of competence and limited validity because only face validity was assessed.
Takase and Teraoka (2011). Development of Holistic Nursing Competence Scale	Competence was conceptualized through concept analysis, in terms of nurses' knowledge, values, professionalism, and motivation to provide nursing care.	The scale was found valid (construct validity indices > 0.45) and criterion validity r = 0.363) and reliable	The conceptual definition was somewhat consistent with the Relational Inquiry approach and several items of the scale were pertinent to the Relational Inquiry approach.	**Strengths** Systematic process for scale development and appropriate statistical analysis. **Limitations** Lack of theoretical framework for conceptualization of

	The sample consisted of 331 nurses. The psychometric properties were assessed through construct and criterion validity and Cronbach's alpha.	(α=0.967).		competence.
Hsu and Hsieh (2013) Development and testing of Competence Inventory to measure nursing students' learning outcomes.	Competence was operationalized in terms of ethical practice, clinical skills, continuous learning, biomedical knowledge, caring and critical thinking. The sample consisted of 599 nursing students. The psychometric properties were assessed through principal and exploratory factor analysis, construct validity and Cronbach's alpha.	The scale was valid (CVI= 0.83-1.00) and reliable (α=0.91-0.98).	The conceptual definition was consistent with the Relational Inquiry approach and several scale items were pertinent to the Relational Inquiry approach.	**Strengths** Systematic process for scale development and appropriate statistical analysis. **Limitations** Lack of theoretical framework for conceptualization of competence.
Nilsson et al., (2014) Development and validation of nurses' Professional Competence Scale	Competence was operationalized in terms of nursing care, values based care, teaching, education, legal practice, medical care, and documentation. The sample consisted of 1086 nursing students. The psychometric properties were assessed: known group, face, and	The scale was valid and reliable (α=0.75-0.96)	The conceptual definition was somewhat consistent with the Relational Inquiry and several items of the scale were pertinent to the Relational Inquiry approach.	**Strengths** Systematic process for scale development and appropriate statistical analysis. **Limitations** Lack of theoretical framework for conceptualization of competence.

	construct validity and Cronbach's alpha.			
Studies on Curiosity				
Kashdan et al., (2009) The development and psychometric assessment of the Curiosity and Exploration Inventory II (CEI-II).	Curiosity was conceptualized in three traits: "active information and opportunity seeking behavior, willingness to accept uncertainty in life, and "tolerance of uncertainty" (p. 989). The techniques of confirmatory factor analysis, convergent and discriminant validity, and item response theory analysis were used. The sample of 578 undergraduate students comprised of three subsamples: 311,150, and 119.	The factor analysis generated two main constructs namely, stretching and embracing. In each sample, the internal consistency of the two constructs and the total scales ranged from α=0.75 to 0.89 which indicates good reliability. The CEI-II was also found to be valid.	In CEI-II the construct stretching was defined as "actively seeking opportunities for new information and experiences" (p. 989). The construct embracing was defined as "willingness to embrace the novel, uncertain, and unpredictable nature of everyday life" (p. 989) These definitions are somewhat consistent with the Relational Inquiry approach. The underlying ideology CEI-II was based on determining students' motivation, learning and well-being. A few operationalized items were pertinent to the Relational Inquiry approach, but they were not appropriate to determine curiosity of the nursing students.	**Strengths** Extensive and systematic process for CEI-II development and appropriate statistical analysis. **Limitations** The CEI-II was more robust for determining the stretching construct rather than embracing because there were only two positive outcome variables; mindful awareness and extraversion that strongly correlated to the embracing scale.

Appendix III

Relational Inquiry Capacities Scale

The Relational Inquiry Capacities Scale (RICS)							
	Items	**Strongly Agree**	**Agree**	**Slightly Agree**	**Slightly Disagree**	**Disagree**	**Strongly Disagree**
		6	**5**	**4**	**3**	**2**	**1**
C1	**Compassion**						
1	I consciously observe patients' emotions during each visit.						
2	I consciously consider patients' feelings, even if they are not expressed.						
3	I attentively listen to patients when they talk to me (extracted and adapted from Pommier, 2010).						
4	I do not impose my own ideas while talking to patients.						
5	I honour what patients share with me.						
6	I inquire about patients' concerns.						
7	I recognize my own fears when caring for patients.						
8	I do not try to fix patients' problems.						
9	I try to be caring towards patients in their difficulties (extracted and adapted from Pommier, 2010).						
10	I feel like I can relate to someone who is feeling down (extracted and adapted from Pommier, 2010).						
11	It's important to recognize that all people have weaknesses and no one's perfect (extracted and adapted from Pommier, 2010).						
12	I like to be with patients in their difficulties.						
13	I like to be with patients in their happiness.						
14	I respect the dignity of patients.						
15	I relate to other nurses and healthcare providers involved in any given nursing situation.						

16	I feel that suffering is just a part of the common human experience (extracted and adapted from Pommier, 2010).						
	Self-Compassion						
1	I consciously examine my own feelings when talking to patients						
2	I examine my own understandings of patients' situations.						
3	I recognize my negative emotions when listening to patients.						
4	I recognize my positive emotions when listening to patients.						
5	I give the same attention to myself as I give to patients.						
6	I care for myself as I care for patients.						
7	I try to see my failings as part of human nature (extracted from Neff's [2003] Self-Compassion Scale)						
8	I acknowledge any of my pre-conceived biases which could affect patient care.						
9	I am approving and non-judgemental about my own limitations in providing patient care (extracted and adapted from Neff's [2003] Self-Compassion Scale).						
10	I often reflect on my caring practices with patients.						
11	When I'm going through a very hard time, I give myself the caring and tenderness I need (extracted from Neff's [2003] Self-Compassion Scale).						
12	I'm kind to myself when I'm experiencing suffering (extracted and adapted from Neff's [2003] Self-Compassion Scale).						
13	I try to be understanding towards those aspects of my personality I don't like (extracted from Neff's [2003] Self-Compassion Scale).						
C2	**Curious**						
1	I consciously seek as much information as I can about any given nursing situation (extracted and adapted from Kashdan's et al., [2009] Curiosity and Exploration Inventory).						
2	I view any given nursing situation as an opportunity to improve						

	my nursing practice (extracted and adapted from Kashdan's et al., [2009] Curiosity and Exploration Inventory).						
3	While caring for patients, I often question my nursing knowledge.						
4	I acknowledge what I do not know when caring for patients.						
5	I use my nursing knowledge when caring for patients.						
6	I understand that nursing knowledge could be uncertain.						
7	I acknowledge that nursing practice is uncertain and unpredictable.						
8	I consciously examine my knowledge about patients' needs.						
9	I am always willing to learn new information about patients.						
C3	**Commitment**						
1	I consciously examine my own nursing values in any given nursing situation.						
2	I do not overlook my own needs while caring for patients.						
3	I do not overburden myself while caring for patients.						
4	I anticipate my personal biases which could affect my interaction with patients.						
5	I try to recognize my nursing commitments in any given nursing situation.						
6	I always try to take responsibility for my nursing actions.						
7	I try to fulfil my commitments with patients.						
8	I always try to fulfil my commitments with patients' families.						
9	I acknowledge my limitations when I fail to meet my commitments.						
10	I acknowledge any cultural differences between the patients and myself which could influence my nursing actions.						
C4	**Competence**						
1	I care for patients without causing any harms.						
2	I consciously evaluate my nursing knowledge during patient care.						

3	I consciously evaluate the needs of my patients (extracted and adapted from Safadi et al., [2010] Competence Scale).						
4	I consciously evaluate my nursing abilities during patient care.						
5	I consciously choose my nursing actions in any given situation.						
6	I recognize any and all factors which may influence patient care.						
7	I do not abuse my power to modify my nursing actions.						
8	I respond to my patients without bias (extracted from Takase and Teraoka's [2011] Holistic Nursing Competence Scale).						
9	I carefully consider patients' choices when planning nursing actions.						
10	I consciously seek answers to any questions arising from my nursing practice (extracted and adapted from Takase and Teraoka's [2011] Holistic Nursing Competence Scale).						
11	I guide my decisions about patient care by taking into consideration the ethical values (extracted and adapted from Meretoja's et al., [2004] Nurses' Competence Scale).						
12	I consciously evaluate my own nursing philosophy before planning patient care (extracted and adapted from Meretoja's et al., [2004] Nurses' Competence Scale).						
13	I prioritize my nursing actions according to changing situations (extracted and adapted from Meretoja's et al., [2004] Nurses' Competence Scale).						
14	I incorporate relevant nursing knowledge to provide optimal nursing care to my patients (extracted and adapted from Meretoja's et al., [2004] Nurses' Competence Scale).						
15	I consciously evaluate patient care outcomes (extracted and adapted from Meretoja's et al., [2004] Nurses' Competence Scale).						
16	I show enthusiasm in carrying out my nursing obligations (extracted and adapted from Safadi et al., (2010) Competence Scale).						

C5	Correspondence						
1	I use a caring attitude while providing care to patients.						
2	I consider the patient’s health status before any nursing action.						
3	I respect patients' opinion when planning nursing care.						
4	I do not perform nursing actions in a hurried manner.						
5	I always think before performing any nursing action.						
6	I do not compare my own nursing actions with other nurses.						
7	I pay attention to the concerns of patients.						
8	I examine the energy I exert in every nursing action.						
9	I consciously observe patients’ surroundings when performing nursing care.						

Appendix IV

Interdisciplinary Committee on Ethics in Human Research

Interdisciplinary Committee on Ethics in Human Research (ICEHR)

St. John's, NL Canada A1C 5S7

Tel: 709 864-2561 icehr@mun.ca
www.mun.ca/research/ethics/humans/icehr

ICEHR Number:	**20180427-NU**
Approval Period:	July 28, 2017 – July 31, 2018
Funding Source:	Not Funded
Responsible Faculty:	Dr. Caroline Porr School of Nursing
Title of Project:	*Measuring Relational Capacities of Nursing Students: Psychometric Testing of the Relational Inquiry Capacities Scale (RICS)*

July 28, 2017

Mr. Ahtisham Younas

School of Nursing

Memorial University of Newfoundland

Dear Mr. Younas:

Thank you for your correspondence of July 18 and 27, 2017 addressing the issues raised by the Interdisciplinary Committee on Ethics in Human Research (ICEHR) concerning the above-named research project.

ICEHR has re-examined the proposal with the clarification and revisions submitted, and is satisfied that the concerns raised by the Committee have been adequately addressed. In accordance with the *Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans (TCPS2)*, the project has been granted *full ethics clearance* to July 31, 2018. ICEHR approval applies to the ethical acceptability of the research, as per Article 6.3 of the *TCPS2*. Researchers are responsible for adherence to any other relevant University policies and/or funded or non-funded agreements that may be associated with the project.

If you need to make changes during the project, which may raise ethical concerns, please submit an amendment request with a description of these changes for the Committee's consideration. In addition, the *TCPS2* requires that you submit an annual update to ICEHR before July 31, 2018. If you plan to continue the project, you need to request renewal of your ethics clearance, and include a brief summary

on the progress of your research. When the project no longer involves contact with human participants, is completed and/or terminated, you are required to provide the annual update with a final brief summary, and your file will be closed.

Annual updates and amendment requests can be submitted from your Researcher Portal account by clicking the *Applications: Post-Review* link on your Portal homepage.

We wish you success with your research.

Yours sincerely,

Kelly Blidook, Ph.D.

Vice-Chair, Interdisciplinary Committee on

Ethics in Human Research

KB/lw

cc: Supervisor – Dr. Caroline Porr, School of Nursing