EFFECTIVE TEACHING IN CLINICAL SIMULATION:

DEVELOPMENT OF THE STUDENT PERCEPTION OF EFFECTIVE TEACHING IN CLINICAL SIMULATION SCALE

Cynthia E. Reese

Submitted in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the School of Nursing,
Indiana University

May 2009

Accepted by the Faculty of Indiana University, in partial
fulfillment of the requirements for the degree of Doctor of Philosophy.

Pamela R. Jeffries, DNS, RN, Chair

Daniel J. Pesut, PhD, RN

Doctoral Committee

Judith A. Halstead, DNS, RN

March 12, 2009

Tamilyn Bakas, DNS, RN

DEDICATION

To: Rick, James and Katie

ACKNOWLEDGEMENTS

I am pleased to express my gratitude to the many people who provided encouragement and support throughout the process of writing of my dissertation. First, I want to acknowledge my committee for their guidance and unwavering belief in my ability to navigate this journey successfully. I would like to thank my chairperson, Dr. Pamela Jeffries for introducing me to clinical simulation. Dr. Jeffries shared her expert knowledge of simulation research and made great efforts on my behalf to ensure that I connected with others in simulation. In addition, Dr. Jeffries ensured that I was welcomed into the wonderful clinical simulation laboratory at IUSON, where my research was conducted. I will be forever grateful for the time, support and mentoring of Dr. Jeffries. I would also like to thank Dr. Daniel Pesut for inspiring and supporting me throughout the doctoral program. Dr. Pesut helped me to better understand myself through the identification of my red thread during my first course in the program, and through the many strategy sessions we had early in my studies. The value of Dr. Pesut's support and encouragement cannot be overstated. Dr. Tamilyn Bakas contributed a great amount of her time and expertise to assist me with the instrument development and measurement aspects of my project. Dr. Bakas gave me the amount of support I needed to be successful and instilled an enthusiasm for instrument development that I will take into future endeavors. I will be forever grateful for the guidance provided by Dr. Bakas. Dr. Judith Halstead helped me to better integrate and understand the role of theory in my writings. Dr. Halstead willingly shared of her time and expertise to ensure that my writings were of the highest quality. Each of the members of my committee played a critical role in my development as a scholar and the successful completion of my dissertation.

Next, I would like to express my gratitude to Shirley Woolf and Lesley Milgrom. These wonderful master teachers allowed me unrestricted access to their course, students and simulations. Without the dedication of Shirley and Lesley and their belief in my project it would have been very difficult to complete my research. In addition, I would like to thank the staff at the IUSON simulation center for their patience and assistance during the data collection for my study.

I am pleased to thank my many friends and colleagues at Lincoln Land Community College for their support and encouragement throughout my journey through the doctoral program and dissertation process. I would like to especially thank my friend, Cynthia Maskey who has been with me from the beginning and her willingness to listen and her unwavering belief in me have meant a great deal.

This dissertation is dedicated to my family. I would like to thank my husband, Rick who has supported me throughout my doctoral studies and all of my educational and professional endeavors. To my daughter, Katie, thank you for your willingness to tolerate my schedule and for your encouragement. I am grateful for all of the sacrifices you have made which allowed me to finish and obtain my degree.

ABSTRACT

Cynthia E. Reese

EFFECTIVE TEACHING IN CLINICAL SIMULATION:

DEVELOPMENT OF THE STUDENT PERCEPTION OF EFFECTIVE TEACHING IN CLINICAL SIMULATION SCALE

Clinical simulation is an innovative teaching/learning strategy that supports the efforts of educators to prepare students for practice. Despite the positive implications of clinical simulations in nursing education, no empirical evidence exists to inform effective teaching in simulated learning environments. The purpose of this research is to create an instrument to measure effective teaching strategies in clinical simulation contexts. The conceptual framework for this study is the Nursing Education Simulation Framework.

The Student Perception of Effective Teaching in Clinical Simulation (SPETCS) is a survey instrument scored on a 5-point Likert scale with two response scales: Extent and Importance. The Extent response scale measures participants' perception of the extent to which the instructor used a particular teaching strategy during the simulation, and the Importance response scale measures perception of the degree of importance of the teaching strategy toward meeting simulation learning outcomes.

A descriptive, quantitative, cross-sectional design was used. Evidence to support content validity was obtained via a panel of simulation experts ($n = 7$) which yielded a content validity index of .91. A convenience sample of undergraduate nursing students ($n = 121$) was used for psychometric analysis. Internal consistency reliability met hypothesized expectations for the Extent ($\alpha = .95$) and Importance ($\alpha = .96$) response scales. Temporal stability reliability results were mixed; correlations between administration times met expectations on the Importance scale (ICC = .67), but were lower than expected on the Extent scale (ICC = .52). Both response scales correlated within hypothesized parameters with two criterion instruments ($p < .01$). The Importance scale was selected for exploratory factor analysis (EFA). EFA revealed 2 factors: Learner Support and Real-World Application. The result of careful item and factor analysis was an easy to administer 33 item scale with 2 response scales.

The SPETCS has evidence of reliability and validity and can serve as a tool for the assessment, evaluation, and feedback in the ongoing professional development of nurse educators who use clinical simulations in the teaching/learning process. In addition, results of this study can support the identification of best practices and teaching competencies in the clinical simulation environment.

Pamela R. Jeffries, DNS, RN, Chair

TABLE OF CONTENTS

LIST OF TABLES x

LIST OF FIGURES xii

CHAPTERS

1. INTRODUCTION 1

Background and Significance 4
Problem Statement 9
Purpose 9
Specific Aims and Hypothesis 10
Conceptual and Operational Definitions 12
Assumptions 15
Limitations 15

2. REVIEW OF THE LITERATURE 17

Overview of the Nursing Education Simulation Framework 17
Socio-cultural Learning Theory 18
Learner-centered Approaches 20
Constructivism 21
The Nursing Education Simulation Framework 24
Teaching Effectiveness 36
Clinical 40
Classroom 50

3. METHODS 61

Design 61
Sample 67
Procedure 68
Human Subjects Approval 69
Variables and Instruments 70
Data Analysis 72
Data Screening Procedures 72
Specific Aims and Hypotheses 73

4. RESULTS 83

Sample 83
Data Screening 88
Specific Aims and Hypotheses 90

5. DISCUSSION 118

Specific Aims 118
Theoretical Implications 125
Research Implications 128
Practice Implications for Nurse Educators 130
Limitations 131
Recommendations for Future Research 133

APPENDICES

Appendix A: Student Perception of Effective Teaching 136

in Clinical Simulation Scale

Appendix B: Recruitment Letter to Content Experts 140

Appendix C: Content Validity Grid 143

Appendix D: Permission to use instruments 149

Appendix E: Nursing Clinical Teaching Effectiveness Inventory 151

Appendix F: Student's Evaluation of Educational Quality 154

Appendix G: 33 Item SPETCS 157

Appendix H: Institutional Review Board Approval 161

Appendix I: Informed Consent Statement 163

Appendix J: Student and Instructor Demographic Survey 166

REFERENCES 169

CURRICULUM VITAE

LIST OF TABLES

1. Description of Clinical Teaching Instruments 47

2. Summary of Dimensions/Factors Identified from Clinical Teaching Instruments......49

3. Classroom Teaching Instruments...51

4. Summary of Dimensions/Factors Identified from Classroom Teaching Instruments 57

5. Student Gender, Ethnicity, Race, Type of nursing program.......................................86

6. Student Age, Grade Point Average (GPA), Previous Simulation Experience, Years of Experience Working in Healthcare 87

7. Instructor Demographic Characteristics 88

8. Missing Data Percentages by Instrument..89

9. Skewness, Kurtosis, Kolomogorov-Smirnov test values by instrument......................89

10. SPETCS Total Scale Descriptive Statistics 92

11. Summary Item Statistics Extent and Importance Response Scales 92

12. Descriptive Statistics SPETCS Extent Response Scale...93

13. Descriptive Statistics SPETCS Importance Response Scale 94

14. Item Analysis Statistics SPETCS Extent Response Scale...96

15. Item Analysis Statistics SPETCS Importance Response Scale 97

16. Correlations and Comparison of Mean Scores Time 1 and Time 2.............................99

17. Comparison of SPETCS Scores Time 1 and Time 2 99

18. Factor Analysis of Item Pool for SPETCS Importance Response Scale 102

19. Factor Analysis of Item Pool for SPETCS Extent Response Scale 105

20. Factor Analysis of the 33 item SPETCS Importance Response Scale.......................108

21. Reliability of 33 item SPETCS Extent and Importance Response Scales 109
and subscales using Cronbach's Alpha

22. Differences in Mean Scores by Instructor 110

23. Correlation between SPETCS and Criterion Instruments using Pearson's *r* 110

24. Rank Order of Items by Mean Score (SD) on the Extent Response Scale 111

25. Rank Order of Items by Mean Score (SD) on the Importance Response Scale 113

26. Correlations of continuous demographic and situational variables to screen 114
for inclusion into regression equation using Pearson's *r*

27. Screening of Discrete Demographic and Situational Variables for 115
Inclusion in Regression Analysis using MANOVA univariate *F*

28. Constructs Underlying Factors in Importance Response Scale 127

LIST OF FIGURES

1. Nursing Education Simulation Framework 10

1. INTRODUCTION

The nursing shortage, the increased use of technology and high patient acuity demands complex decision making, often under conditions of uncertainty and risk (Simmons, Lanuza, Fonteyn, Hicks & Holm, 2003). Nursing graduates are faced with a health care environment in which they have greater responsibility and accountability than in years past. Consequently, nurse educators must prepare graduates capable of successfully meeting the multifaceted demands of the workplace. Clinical simulation is an innovative teaching/learning strategy that supports the efforts of educators to prepare students for practice (Hotchkiss, Biddle & Fallacaro, 2002; Nehring, Lashley & Ellis, 2002; Bearnson & Wiker, 2005; Cioffi, 2005; Jeffries, 2005; Bradley, 2006). Despite the positive implications of clinical simulations in nursing education, there is a paucity of evidence to support overall best practices, and no evidence to inform effective teaching in simulated learning environments. The development and validation of a psychometrically sound instrument to measure teaching effectiveness in clinical simulations is a significant research challenge that supports the identification of evidence-based best practices in teaching in simulations.

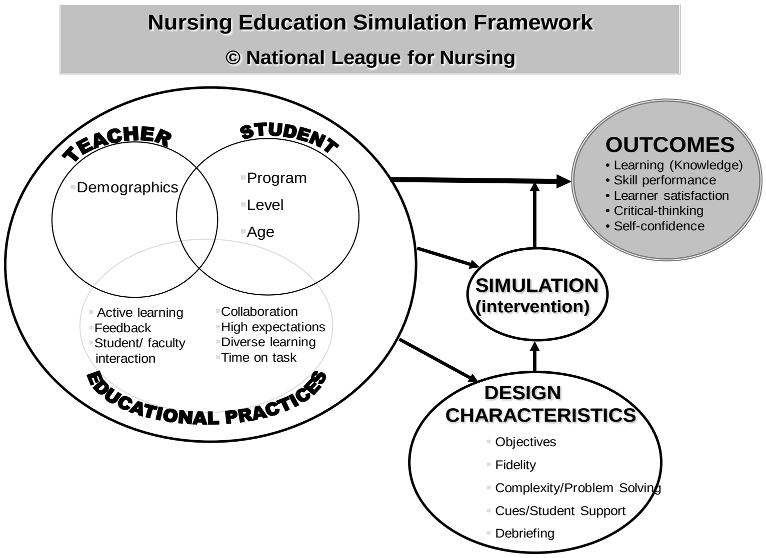
The use of clinical simulation in all health professions education has grown exponentially in recent years and this growth is expected to continue in the future (American Association of Medical Colleges [AAMC], 2000; American Association of Colleges of Nursing [AACN], 2005; National League for Nursing [NLN], 2005). Clinical simulation is an experiential teaching/learning strategy that facilitates students' ability to apply and synthesize knowledge learned in the classroom in an environment that imitates the clinical setting (Jeffries, 2005). Through various types of high-fidelity or realistic simulations, it is hypothesized that students can develop and improve decision-making, clinical judgment, self-confidence, and psychomotor abilities in the safe environment of the skills laboratory (Shearer, Bruce, Graves & Erdley 2003; Cioffi, Purcal & Arundell 2005; Jeffries, 2005; Bradley, 2006). As the demands of the workplace require greater cognitive, affective and psychomotor competencies from graduates, simulations are hypothesized to improve learning outcomes when used to supplement traditional classroom and clinical instruction (Johnson, 1999; Bearnson & Wiker, 2005).

Although the use of clinical simulation as a teaching/learning strategy in health care education has increased, the significant majority of the literature describing simulations is anecdotal. There is little research to date that describes or evaluates the role or behaviors of the educator and the educational outcomes of students working in clinical simulations (Seropian, 2003; Jeffries, 2004; Cioffi, et al. 2005). Most empirical studies in simulation are atheoretical, lack methodological rigor, have small sample sizes, and lack reliable and valid measures (Bradley, 2006). In light of the complexity of the health care system, and the need to apply evidence-based approaches in healthcare education, the development of an evidence base describing and evaluating clinical simulations is essential (Reed, et al., 2005).

To address the need to develop an evidence base supporting the learning outcomes of clinical simulation, and to guide the design, implementation and evaluation of simulation, a conceptual model based on learner-centered, socio-cultural, constructivist learning theory has been developed (Jeffries, 2005). The Nursing Education Simulation Framework (NESF) has five major components: 1) Educational practices, 2) Teacher, 3) Student, 4) Design characteristics and 5) Outcomes (Figure 1). The NESF proposes that the degree to which educational outcomes are achieved depends on the use of best educational practices (Chickering & Gamson, 1987). Best educational practices include both careful attention to the interactions between students and teachers and the use of empirically supported design features as simulations are designed, implemented and evaluated. Each component of the NESF has been tested and evaluated with the exception of variables related to the teacher (Jeffries, 2005; NLN, 2006).

The role of the teacher and teaching effectiveness in clinical simulations are major components of the Nursing Education Simulation Framework that have not been described or tested empirically. There are several instruments measuring teaching effectiveness in the classroom and clinical settings (e.g. Buskist, 2002; Copeland & Hewson, 2000; Leamon & Fields, 2005; Marsh 1987; Mogan & Knox, 1987; Mourad & Redelmeier, 2006). However, no instrument exists to measure effective teaching within the context of clinical simulation. Additionally, the development of measures created for use in specific learning environments allows more accurate and detailed feedback to faculty (Raingruber & Bowles, 2000). Due to the lack of tools to measure the unique educational features of simulations, a new instrument has been created. The purpose of this research study was to create an instrument to measure effective teaching strategies in clinical simulation contexts. The SPETCS can serve as a tool for assessment, evaluation, and feedback in the ongoing professional development of nursing educators who use clinical simulations in the teaching learning process.

Figure 1.



<u>Background and Significance</u>

The nursing faculty shortage, cost containment issues and the unavailability of appropriate clinical sites for students continue to be problematic for many schools of nursing (AACN, 2005; Feingold, Calaluce & Kallen, 2003; Medley & Horne, 2005; NLN, 2005). These challenges have forced many nursing education programs to alter the structure of their curriculum and decrease the number of traditional clinical hours. Clinical simulations have the potential to mitigate the impact of these issues through development of innovative models of clinical education.

The nurse faculty shortage is significant, and likely to worsen in the foreseeable future (NLN, 2005; AACN, 2005). The Bureau of Labor Statistics projects a deficit of one million registered nurses by the year 2010. The AACN (2005) has projected that at least 390,000 of the possible one million vacant positions for registered nurses will be nurses with baccalaureate or masters degrees and as a result, there is an urgent need for added faculty to educate future nurses (AACN, 2005). Compounding the issue is the shortage of master's and doctoral prepared faculty which limits admissions to schools of nursing. Two primary rationales are cited for the faculty shortage; the aging of faculty and a decreasing population of younger faculty (AACN, 2005; NLN, 2005). Consequently, the demand for future faculty is greater than the supply and in turn, this directly influences the numbers of nurses that can be educated at all levels.

The use of supervised simulated clinical experiences in learning resource centers is beginning to replace some of the clinical hours to reduce faculty workload. Nehring, Lashley & Ellis (2002) and Mc Causland, Curran & Caltaldi (2004) recommend designating one faculty member or alternatively a small group of faculty to train and operate the simulation program. This could potentially decrease the total number of faculty needed when compared to the usual 8:1 or 10:1 ratio of students to faculty required for clinical instruction. A decrease in existing faculty workload by may allow schools of nursing to admit an additional cohort of students. In turn, the increase in enrollment may moderate the nursing shortage.

Cost containment and budgetary constraints continue to require creative solutions to support the educational needs of schools of nursing (AACN, 2005; NLN, 2005). While it is recognized that the initial investment in simulation technology and training is high, with cost estimates ranging between $30,000 to $200,000, these costs may ultimately decrease following the initial purchase and implementation of simulations into the curriculum when compared to the high costs of clinical instruction (Mc Causland, Curran & Caltaldi, 2004; Rauen, 2001; Scherer, 2003). Additionally, monetary support for simulation technologies and research is increasing primarily through federal sources such as the Association of Health Care Quality, and National Council of State Boards of Nursing.

Increased competition for clinical sites among schools of nursing continues to be problematic as enrollments increase and the number of inpatient days of hospital care continues to decrease (Medley & Horne, 2005). As a result, educators need to consider a change in current models of clinical teaching. One emerging model is the replacement of clinical hours with hours in the simulation laboratory. To address the issue of the substitution of clinical hours with hours in clinical simulations, state boards of nursing are beginning to develop guidelines for the percentage of clinical time that may be replaced by simulations. Currently, the states of California, Florida, and Colorado have provisions in their respective nurse practice acts specifying the percentage of clinical hours that may be supplanted by hours in the simulation laboratory (www.rn.ca.gov/meetings/pdf/elc-may06.pdf). These changes in clinical models of teaching highlight the need to ensure that simulations are well-developed, educationally sound, and evidence-based. Overall, simulations have the potential to make a positive impact on several fronts, specifically the faculty shortage, budgetary concerns, and lack of availability of clinical sites.

In addition to the positive impact of clinical simulations on those issues described above, the literature describing clinical simulations in health care professions advances several additional reasons to support increased use of simulation as a teaching/learning and evaluative strategy: 1) technology has improved the fidelity and decreased the cost of available simulators, 2) active learner participation is expected, 3) students are exposed to similar complex problems through standardized experiences,

4) simulated learning provides a low risk environment without risk of harm to patients, 5) simulation enables the development of clinical reasoning and problem solving skills which may be applied in the clinical setting with actual patients, 6) simulations enable immediate feedback from faculty and 7) policy-driven emphasis on improved patient safety and outcomes supports simulated learning (Fletcher, 1995; Aronson, Rosa, Anfinson & Light, 1997; Institute of Medicine, 2003, Feingold, et al., 2004; Spunt, Foster & Adams, 2004; Medley & Horne, 2005; Bradley, 2006).

In contrast, several challenges related to the use of simulation have been described in the literature. These include the: 1) cost of purchasing simulation equipment, 2) large initial faculty time commitment, 3) faculty hesitation to adopt simulation, 4) inappropriateness for large group instruction and 5) lack of evidence validating knowledge transfer and improved patient outcomes in the clinical environment as a result of student exposure to simulation training (Nehring & Lashley, 2004; Scherer et al., 2003; Seropian et al., 2003). These potential issues must be considered when integrating simulation into an educational program.

Effective teaching in clinical simulations has a powerful impact on student experiences and outcomes of a simulation exercise (Nehring, et al. 2002; Henneman & Cunningham, 2005; Bremner, Aduddell, Bennett & VanGeest, 2006). The role of teacher and effective teaching strategies are informed by the learner-centered, socio-cultural, and constructivist learning theories underpinning the NESF. The role of the teacher changes significantly with learner-centered and constructivist teaching as the focus shifts from teacher to learner control of the educational environment. The teacher is viewed as a facilitator of learning who structures learning experiences to allow students to construct knowledge for themselves, in contrast to a lecture driven, content delivery mode of instruction. Additionally, the educational objectives of the simulation direct the teaching strategies necessary for a successful simulation. For example, a simulation developed for teaching purposes requires the instructor to perform in a more supportive and facilitative role; however if the simulation is used to evaluate students, the appropriate role is more of an observer (Jeffries, 2006, 2007). As discussed above, effective teaching strategies in simulation have not been identified. It is hypothesized that through the identification and subsequent use of effective teaching strategies, the quality of simulations will improve, which will ultimately improve the ability of nursing graduates to function safely in the workplace.

Although the use of high-fidelity clinical simulations in nursing education is increasing dramatically, the literature to support simulation has not kept pace and is just beginning to develop (Bradley, 2006; Feingold, et al., 2004; Jeffries, 2005; McCausland et al., 2004). It is no longer considered acceptable to utilize teaching/learning strategies supported by anecdotal literature alone (NLN 2005, Nehring, Ellis & Lashley, 2001). The role of the teacher and use of effective teaching strategies are integral components of successful simulation experiences and the attainment of educational outcomes for students (Jeffries, 2006; Seropian, 2003). Although there is empirical literature describing effective teaching in the classroom and clinical contexts, the anecdotal literature pertaining to effective teaching in simulation is sparse, and empirical studies are non-existent (Bradley, 2006; Jeffries, 2006). In addition, several authors have cautioned against extrapolating empirical results related to effective teaching between teaching contexts, which highlights the need for context-specific research in simulations (Marsh & Roche, 1997).

In light of the growth of simulation as a teaching/learning strategy, the lack of context-specific research related to teaching effectiveness in clinical simulation and the hypothesized importance of the teaching role to the achievement of student outcomes based on the NESF, there is an urgent need for empirical study to address this significant gap in the literature. The development and psychometric testing of an instrument to evaluate teaching effectiveness in simulation environments is a necessary first step to expand the empirical knowledge supporting the effective use of clinical simulation in health professions education.

Problem Statement

The use of clinical simulation in nursing education has grown exponentially in recent years. The role of the teacher and teaching effectiveness in clinical simulations are major components of the NESF which have not been described or tested empirically. Instruments are available to measure teaching effectiveness in the classroom and clinical settings, no instrument exists to measure effective teaching within the unique context of high fidelity clinical simulation.

Purpose

The purpose of this study was to develop and psychometrically test a theory-based instrument to measure effective teaching strategies and behaviors specifically in clinical simulation contexts. The conceptual framework for this study was the NESF, which guides the design, implementation and evaluation of clinical simulations, and is based on learner-centered, socio-cultural and constructivist learning theories

(Jeffries, 2005; 2007). The SPETCS has the potential to serve as a tool for assessment, evaluation, and feedback in the ongoing professional development of nursing educators who use clinical simulations in the teaching learning process.

Specific Aims and Hypotheses

The specific aim of this research was to create an instrument with evidence of psychometric reliability and validity that measures effective teaching in clinical simulation contexts.

Specific Aim 1. Develop the Student Perception of Effective Teaching in Clinical Simulations (SPETCS) scale and determine the degree of relevance of individual items and the overall scale using Lynn's (1986) content validity index (CVI).

Hypothesis 1a. The SPETCS scale items demonstrates a CVI of ≥ .78. Using a panel of nine experts, seven need to be in agreement regarding the relevance of each item to obtain an acceptable CVI of ≥ .78 (Lynn, 1986).

Hypothesis 1b. The SPETCS total scale demonstrates a CVI of ≥ .78, which is the proportion of total items judged content valid (Lynn, 1986).

Specific Aim 2. Evaluate the psychometric properties of the SPETCS scale.

Hypothesis 2a. The SPETCS scale items demonstrate means near the center of the scale (3), the range of standard deviations indicates variability in the data, and floor and ceiling effects are less than 10% (DeVellis, 2003).

Hypothesis 2b. Item-to-total correlations for items in the SPETCS scale are ≥.30 among student participants in simulation (Ferketich, 1991).

Hypothesis 2c. The SPETCS scale demonstrates evidence of internal consistency reliability with Cronbach's alpha values ≥.70 among student participants in simulation (Carmines & Zeller 1979; DeVellis, 2003; Nunnally, 1978)

Hypothesis 2d. Evidence of temporal stability of the SPETCS scale is provided by 2 week test-retest reliability with an intra-class correlation coefficient > .60 among student participants in simulation (Shrout & Fleiss, 1979; Yen & Lo, 2002).

Hypothesis 2e. The dimensionality and initial evidence of construct validity of the SPETCS are provided with factor loadings of .32 (Tabachnik & Fidell, 2001), and above for each domain as determined through exploratory factor analysis using principle axis factoring among student participants in simulation (Netemeyer, Bearden & Sharma, 2003).

Hypothesis 2f. The SPETCS scale demonstrates evidence of criterion-related validity as evidenced by significant ($p < .05$) correlation with the Student Evaluation of Educational Quality (SEEQ, Marsh, 1987) and the Nursing Clinical Teacher Effectiveness Instrument (NCTEI; Mogan & Knox, 1987).

Specific Aim 3. Determine which teaching strategies/behaviors are most frequently used in clinical simulations based on perceptions of student participants.

Specific Aim 4. Determine which teaching strategies/behaviors best facilitate achievement of specified simulation outcomes based on ratings of student participants in simulation.

Specific Aim 5. Determine the relationship between student participant's demographic and situational variables and their responses on the Extent and Importance response scales of the SPETCS.

Hypothesis 5a. There are no significant differences between students' mean scores on the Extent and Importance response scales between student groups who have different instructors facilitating the simulation experience.

Conceptual and Operational Definitions

Simulation

Conceptual Definition. "Simulations are highly realistic operating models that contain relevant central features of reality which function over time. Simulations are designed to represent, model, clarify, explain, predict, and/or communicate the features, impacts, and intrinsic interrelationships of the modeled system of interest. When persons participate, the task and task environment must be perceived as sufficiently real to produce relevant real-world equivalent (in contrast to imagined or role played) behavior" (Streufert, Satish & Barach, 2001, p. 167).

Human Patient Simulator (HPS)

Conceptual Definition. "The human patient simulator is a life-like manikin with sophisticated computer controls that can be manipulated to provide various physiological parameter outputs being either physical or electrical in nature or a combination of the two. The parameters may be controlled through automated software or respond to the actions of an instructor in response to student actions" (Bradley, 2006, p. 258).

Fidelity

Conceptual definition. Fidelity is the degree of realism or accuracy of the system used, and simulations may be categorized as low, medium or high-fidelity (Seropian, 2004). Sophisticated human patient simulators which are designed to closely mimic reality are considered to have a high degree of fidelity.

Teaching Effectiveness in Clinical Simulation

Conceptual definition. Effective teaching in clinical simulation is the degree to which the teaching strategies and behavioral characteristics of the instructor promote student achievement of the learning outcomes specified in the simulation experience.

Operational definition. Effective teaching was assessed by the SPETCS. The SPETCS contained two response scales, Extent of Agreement and Importance. The Extent response scale represented students' extent of agreement with items related to specific teaching strategies and behaviors used during the simulation. The Importance response scale represented students' perception of the degree of importance the teacher behavior or strategy in each item was toward meeting the specified outcomes of the simulation. Each scale was scored on a 5 point Likert scale. The Extent response scale asked students to indicate their extent of agreement with each item depicting a teaching behavior or strategy used during the simulation: SD) strongly disagree, D) disagree, N) neutral (neither agree nor disagree), A) agree and SA) strongly agree. The Importance response scale was scored based on student perceptions of how important the teaching behavior or strategy used during the simulation was to meeting the simulation learning outcomes: 1) not important, 2) slightly important, 3) moderately important, 4) very important and 5) extremely important.

Individual item scores and summed scores were analyzed. Individual items with the highest mean scores in the Extent response scale were those behaviors or strategies most frequently used in the simulation. The items with the highest means on the Importance response scale represented student participants' perceptions of which instructor behaviors or strategies were most important to meet the specific simulation objectives. Additionally, summed subscale scores were examined with higher total scores on the Extent response scale indicative of more frequent use of teaching strategies and behaviors. Summed scores on the Importance response scale were indicative of the importance of all of the teacher behaviors and strategies towards meeting simulation objectives.

Demographic and Situational Variables

Conceptual definition. Demographic and situational variables which may affect the study results were measured from student participants to assess the influence or potential relationship of these variables to the results of the SPETCS. Additionally, the demographic data was collected on the two instructors involved with the simulation to provide a clear description of the study sample.

Operational definition. A demographic data form measured student participants' age, gender, ethnicity, race, grade point average, accelerated or traditional track student, amount of experience with simulations, and work experience in healthcare settings. Situational variables such as level in undergraduate program, university, and course were identical for all students. Participants were in two groups, each with a different instructor. Demographic data collected from the instructors included age, sex, ethnicity, race, certification status, educational preparation, number of years of teaching experience and the amount of experience teaching in simulation contexts.

Simulation Expert

Conceptual definition. Nursing faculty who are designated by the NLN as simulation experts were considered as experts for the content validation of survey items for this study.

Nursing Faculty with Simulation Experience

Conceptual definition. Faculty with simulation experience are those from all levels of nursing education, undergraduate through graduate who participate in the development, implementation, and evaluation components clinical simulation.

Assumptions

1. The SPETCS accurately measures the construct of effective teaching in clinical simulation.

2. Effective teaching behaviors and strategies can be measured using a self-report survey instrument.

3. Students are able to assess faculty use of effective teaching behaviors and strategies in clinical simulations.

4. Student evaluations of teaching serve as a tool for assessment, evaluation and feedback in the ongoing professional development of nursing educators who use clinical simulations in the teaching learning process.

Limitations

1. A non-probability, homogeneous, convenience sample was used in this study.
2. Characteristics of effective teaching were measured by a newly developed instrument with no prior psychometric analysis.
3. A cross-sectional, descriptive design with a two week retest was used.

The three limitations outlined for this study were considered to be acceptable due to the descriptive nature and purpose of the research. A homogeneous sample was selected for psychometric analysis of the new instrument. To date, no instrument or evidence exists to measure effective teaching in clinical simulations. The empirical measure developed and validated through this research has the potential to guide future research into teaching effectiveness and support the professional development of faculty utilizing clinical simulation as a teaching/learning strategy in nursing education.

Overview of Chapters

There are 5 chapters in this dissertation. Chapter 1 introduces the problem, describes the background and significance of the problem, defines key terms pertinent to the study, and presents the research purposes, questions, assumptions, and limitations. Chapter 2 is a review of the literature relevant to the purposes of the study. Literature in the areas of clinical simulation and teaching effectiveness in the classroom and clinical areas are reviewed. Chapter 3 is a detailed description of the research design and methodology used to collect and analyze the data. Chapter 4 presents the results of the comprehensive psychometric analysis outlined in Chapter 3. Chapter 5 discusses the implications of the findings in this study

2. REVIEW OF THE LITERATURE

As discussed in the introduction, the literature describing effective teaching in simulations is sparse, as the literature set consists of anecdotal accounts and lacks empirical evidence. The Nursing Education Simulation Framework (NESF) provides guidance for the development, implementation, and evaluation of clinical simulations in nursing; however the characteristics of effective teachers in the model have yet to be examined (Jeffries, 2006).

This chapter provides an in depth review of the NESF and the learning theories supporting the model. In addition, the research related to the measurement of effective teaching in the classroom and clinical areas pertinent to this study is reviewed. A synthesis and critique of the related educational literature follows as a basis for development of the Student Perception of Effective Teaching in Clinical Simulation scale (SPETCS).

Overview of the Nursing Education Simulation Framework

Given the need to develop empirical support for the use of clinical simulation in nursing, a simulation model (Jeffries, 2005) has been developed to guide the design, implementation, and evaluation of simulations. The NESF provides a framework grounded in a synthesis of learner-centered, socio-cultural, and constructivist learning theories. Chickering and Gamson’s (1987) best practices in undergraduate education provided a framework for the development of the educational practices component of the model. The following discussion describes the theoretical underpinnings of the NESF.

Social-cultural Learning Theory

Social constructivism, the educational theory supporting the Nursing Education Simulation Framework, was developed by Vygotsky (1978), and originated from Marxist theory. Vygotsky rejected the reductionist psychology of his day, arguing that to understand an individual, one must first understand the context in which the individual exists (Lattuca, 2002). Learning, in social constructivist terms is “the development of connections with and appropriation from the socio-cultural context in which we all exist” (Bonk & Cunningham, 1998, p. 33). Palincsar (1998) describes learning and understanding as “inherently social, with cultural activities and tools such as language and mathematical symbols, are regarded as integral to conceptual development” (p. 348). As such, the examination of the educational context and setting is essential in order to understand and facilitate student thinking and learning.

Social constructivist teaching emphasizes active learning methods embedded within an environment in which the role of educator is that of a facilitator or guide. Discussions and interactions between learners as well as between the educator and learners help to transform the learning context into a learner-centered environment. Through these interactions, the educator can model his or her reasoning processes which serve as a scaffold for the development of student knowledge (Lattuca, 2006). Scaffolding is a temporary framework developed by the educator which supports the development of the learner and is gradually withdrawn as the learner develops understanding (Dunphy & Dunphy, 2003; Sanders & Welk, 2005). Scaffolding supports the student initially, and the support is gradually withdrawn as the student actively constructs understanding in a way that makes meaning for them (Sudzina, 1997).

A central tenet of socio-cultural learning theory is the concept of the zone of proximal development (ZPD). Vygotsky (1978) defines the ZPD as “the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” (p. 85). Thus, productive interactions between peers and educators become a critical element to learn and develop higher order thinking and problem solving abilities. Additionally, it is hypothesized that interactions between the learner and a more capable peer becomes more important than those between educator and learner and should be encouraged (Palincsar, 1998).

Strategies to scaffold student learning, known as assisted performance and based on the application of the ZPD, have been created to apply socio-cultural learning theory into medical and nursing education (Dunphy & Dunphy, 2003; Sanders & Welk, 2005). Five strategies to assist student performance have been identified as modeling, feedback, instructing, questioning, and cognitive structuring. As the student progresses in knowledge and ability, the scaffolding supports are slowly removed and the student begins to internalize information and perform without assistance. The responsibility for learning shifts during this time from the teacher to the learner, and it is hypothesized that teaching occurs when assistance is offered at certain points in the student’s ZPD where assistance is required (Dunphy & Dunphy, 2003; Vygotsky, 1978).

In addition to assisting individual student performance, learning activities reflecting the complexity of real-world problems allow students to make meaning and develop a deeper understanding of realistic situations. Socio-cultural learning theory implies that learning best occurs within realistic environments, which is one of the primary advantages of simulations over didactic learning contexts. Thus, the basic tenets of social constructivism fit well with simulation as a teaching/learning strategy, the NESF, and the role of educator in clinical simulation.

Learner-centered Approaches

Weimer (2002) describes teaching in a learner-centered environment as a context in which teachers “position themselves alongside the learner and keep the attention, focus, and spotlight aimed at and on the learning process” (p. 76). The learner-centered role of teacher generally defined as a facilitator, coach, or guide. Further, Weimer (2002) proposes five key areas for educators to examine to facilitate changes in practice to move towards a learner-centered environment (pp. 8-17):

1. The balance of power. Power must shift from the teacher to the learner. The focus of the classroom must be on learners.
2. The function of content. This must be taken into consideration as rote memorization of information leads to “surface learning” as opposed to “deep learning.” Deep learning involves relating previous knowledge and what has been learned from different courses to new knowledge. The knowledge is internally analyzed and integrated by the learners into a coherent whole.
3. The role of the teacher. The teaching role is one of a facilitator of learning rather than a transmitter of knowledge seen in traditional teaching.
4. Responsibility for learning. Students bear much of the responsibility for their own learning, as the focus is clearly on the learner. This may pose a challenge to students initially as traditionally educators take responsibility

for organizing content and information for students to learn. Increasing student responsibility for learning may encourage students to develop as autonomous learners.

5. Evaluation purpose and process. Traditional classroom assessments are driven by grades using assessment and evaluation tools developed by the educator. The learner-centered classroom supports the use of evaluation activities which enhances learning, and provides opportunities to develop self and peer-assessment skills.

The use of learner-centered principles in the design, implementation, and evaluation of clinical simulation may improve student outcomes outlined by the NESF. The empowerment of students in a learner-centered environment encourages the growth of students into confident, self- motivated learners.

Constructivism

Constructivism as a theory of learning has developed past its philosophical roots and has garnered support from the fields of neurobiology and cognitive science (Fosnot, 2005, .p.276). A consensus among educational researchers has developed with agreement that knowledge is actively constructed rather than transmitted to the learner. Von Glaserfeld (1996) describes two basic aspects of the constructivist model: 1) learning is a constructive activity that the students themselves must carry out and 2) the task of the educator is not to dispense knowledge but to provide students with opportunities and incentives to build it up. However, what continues to remain challenging for educators are the implications for the application of constructivist theory to education, specifically the implications of constructivism in regards to teaching strategies.

Capra (2002) developed a new biological model which informs constructivism. Cells and cellular networks are described as complex, nonlinear, open systems. Whole cellular networks are thought to exist, and these living and evolving systems need a continual flow of matter and energy from the environment to stay alive. An integral facet of this model hypothesizes that when a flow of energy to the system increases, the overall system develops instability which results in the development of a bifurcation point. The bifurcation point is the location that produces new structures which developed as a direct result of interaction with the environment. As such, an individual's connection with the environment triggers changes in neural structure. These changes have direct implications for the construction of knowledge based on prior understandings and challenges posed to the learner from the environment (Fosnot, 2005 p. 278).

Recent work in cognitive science is lending support to constructivism. Learning is more complex than previously thought and older linear models were found to be inadequate. Cognitive scientists have begun to think about learning and behavior as determined by the structure of an organism and these structures are continuously changing as an individual interacts with the environment. Thus, when the individual interacts with the environment and their current understanding of the world becomes insufficient, when learning a new concept, the bifurcations mentioned above result and new structures are developed. New information is interpreted and organized based on cognitive structures previously developed (Fosnot, 2005). Ultimately, the result is the development of new mental images allowing the formulation of values, beliefs, goals, and strategies. Those ideas are represented through the use of language and mathematical symbols defined and interpreted within social communities of inquiry and discourse. The result of this process, according to cognitive scientists, is epistemological evolution or learning.

As support for the epistemological foundations of constructivism continues to develop, challenges remain regarding the application of the theory into the educational setting. Constructivism is a theory of learning and not a theory of teaching. However, based on an understanding of constructivism, an educator is able to develop applicable teaching strategies. In a global sense, Fosnot (2005) describes the role of the teacher in the constructivist tradition as a facilitator, provocateur, and questioner focusing discussions around big ideas and efficient strategies aligned with learning. Richardson (2003) outlines the following characteristics of constructivist pedagogy:

1. Attention to the individual and respect for students' background and developing understandings and beliefs about elements of the domain.
2. Facilitation of group dialogue that explores and element of the domain with the purpose of leading to the creation and shared understanding of a topic.
3. Planned and often unplanned introduction of formal domain knowledge into the conversation through direct instruction, reference to text, exploration of a web site, or some other means.
4. Provision of opportunities for students to determine, challenge, change, or add to existing beliefs and understandings through engagement in tasks that are structured for this purpose.
5. Development of students' metawareness of their own understandings and learning processes.

While the characteristics of constructivist teaching have been explicated by Richardson (2003), behaviors corresponding to those characteristics have not been developed. However, these global characteristics of constructivist pedagogy may be relevant to the identification of effective teaching behaviors within clinical simulation. For example, specific behaviors relating to the facilitation of group dialogue during the debriefing period following simulation have been suggested. These behaviors include: 1) structuring the discussion using preset questions, 2) seating the group in a circle, 3) maintaining eye contact, 4) focusing on the positive aspects of the simulation and 5) using a non-judgmental communication style. Thus, the lack of development of specific theory-based teaching behaviors is problematic for clinical simulation design, implementation, and evaluation.

A synthesis of the theories and principles underpinning the NESF: socio-cultural learning theory (Vygotsky, 1978), the principles of learner-centered teaching (Weimer, 2002), constructivist characteristics of teaching (Fosnot, 2005; Richardson, 2003) and, Chickering and Gamson’s (1987) best practices provides a set of principles to guide educators working with simulation as a teaching-learning strategy. However from a practical standpoint, using guiding principles that are not explicit, defined, or measureable are of limited usefulness to front-line educators. There is a pressing need to identify effective teaching behaviors in simulation to enhance, coach, and evaluate teaching within this context.

The Nursing Education Simulation Framework

In 2006, a multi-year, multi-site trial supported by grants from both corporate and non-profit groups to study simulation in nursing education was completed. The simulations were developed using the NESF as the conceptual framework, and the research was standardized using instruments specifically developed to test the model (Jeffries & Rizzolo, 2006). The NESF (Figure 1, p. 4) supports the development of simulations and standardizes research examining the use of simulation in nursing education and practice. The results provided empirical evidence supporting the simulation design features, educational practices and educational outcomes identified in the model. However following a review of the literature, a significant gap was found in regard to the role of the teacher and effective teaching behaviors in clinical simulation. The following discussion describes the relationships between the concepts described in the NESF.

Teacher Factors

Effective teaching in clinical simulations has a strong impact on student experiences and outcomes of a simulation exercise (Nehring, et al. 2002; Henneman & Cunningham, 2005; Bremner, Aduddell, Bennett & VanGeest, 2006). The role of teacher and effective teaching behaviors are informed by the learner-centered, socio-cultural, and constructivist learning theories which guided the development of the NESF. The role of the teacher changes significantly with learner-centered and constructivist teaching as the focus is less on teaching and more on learning. The role of the teacher and associated behaviors depends in large part on the objectives of the simulation. A simulation utilized for teaching purposes requires the teacher to perform in a more supportive and facilitative role; however if the simulation is used to evaluate students, the teaching role becomes more of an observer (Jeffries, 2006). As discussed above, effective teaching behaviors in simulation have not been identified. As a result, this research identifies effective teaching behaviors. It is hypothesized that through the identification and use of effective teaching behaviors in clinical simulation, the achievement of expected student outcomes may improve.

Student Factors

According to the NESF, student factors need to be taken into account when designing a simulation; these factors include student age, educational program, and educational level within their program. Generally, beginning students require support and encouragement as they may be unfamiliar with the technological aspects of the simulator; they may possess a more limited knowledge base and higher anxiety level than more advanced students. However, younger students may be more comfortable with technology than older students. The theoretical underpinnings of the NESF support the need to assess learners prior to simulation to ensure that the simulation is developed and designed appropriately in order to promote learning and achieve outcomes.

Expectations of students in simulation can be informed by the learning theories guiding the NESF. Students are expected be responsible for learning, with increasing responsibility as the student progresses through the educational program (Fosnot, 2005; Richardson, 2003; Weimer, 2002). Students need to be motivated and self-directed. Clear objectives for the simulation may allow students to be self-directed to a greater degree. Jeffries (2006) suggests that competition between students may inadvertently occur in simulation, but should be discouraged to decrease learner anxiety and stress. The reduction of learner anxiety and stress in the learning environment allows optimal learning to occur. The role of the student and student factors in simulations has not been described or studied empirically, and is an area for future research.

Educational Practices

Educational practices identified in the NESF are based on Chickering and Gamson's Seven principles for Good Practice in Undergraduate Education (1987). The best practices were identified by a panel of experts in educational research and are based on a synthesis of evidence describing excellence in teaching and learning. These principles are applicable to any college or university undergraduate program (Chickering & Gamson, 1999). Many individual faculty, colleges and universities have adapted these principles in course and program development, faculty professional development and research into student experiences at undergraduate and doctoral level universities (Chickering & Gamson, 1999; Kuh, Pace & Vesper, 1997). Additionally, the principles have been adapted for use in various teaching-learning contexts (Chickering & Ehrmann, 1996). Good practice in undergraduate education includes: active learning, feedback, student/faculty interaction, collaboration, time on task, diverse learning and high expectations. Any of these individual practices may stand alone, however when all are used together the effects are synergistic (Chickering & Gamson, 1987).

The NESF suggests a relationship between the teacher and student factors and the educational practices. Simulations designed using the aforementioned practices, while taking teacher and student factors into account may improve the educational outcomes of the intervention. The tenets of the Nursing Education Simulation Framework suggest that the concomitant use of all seven principles when defining the educational practices of a simulation increase student learning and satisfaction (Jeffries, 2005). The following discussion explicates each of the seven principles.

Active Learning

Active learning strategies are those which engage the student into the learning process. Active learning is an important component of learner-centered, constructivist, and socio-cultural learning theories as it is postulated to support student higher order thinking and complex problem solving (Bonk & Cunningham, 1999; Fosnot, 2005; Vygotsky, 1978). By nature, clinical simulations are an active learning strategy allowing students to apply, evaluate, and synthesize their developing knowledge in a way that mimics reality and may be transferrable to the clinical setting (Cioffi, 2001).

Feedback

Good practice encourages feedback prior to the learning situation to help the learner assess their knowledge. Feedback during learning is used to assess performance. Feedback during the debriefing allows students to reflect and process how their prior knowledge influences simulation performance. Reflection before, during and following actions taken during simulations may improve future learner performance and facilitate the integration of theory into clinical practice (Schon, 1987; Greenwood, 2000). In addition, the use of a non-judgmental debriefing style by the educator has been hypothesized to improve student learning with simulations as the instructor focuses on assisting the student to understand their thinking and reframe that thinking to improve future practice. Additionally, students become less anxious and thus more willing to share their thoughts and become more open to feedback from faculty and peers (Rudolph, Simon, Dufresne & Raemer 2006).

Student/Faculty Interaction

Chickering & Gamson (1987) cite frequent interaction between students and faculty as the most important factor in student motivation and involvement. Good communication helps to keep students focused and working when facing a challenge. In simulation contexts, the amount and content of interaction during the simulation exercise is dependent upon the objectives. Simulations used for practice or learning requires greater student/faculty interaction than simulations used for evaluative purposes (Jeffries, 2006).

Collaboration

Working together in a clinical simulation may improve the communication, problem solving and decision making of the student group. Additionally, learning to work together as a team is a core educational competency cited by the Institute of Medicine (2002) as a mechanism to improve safety and patient care. Thus, it is important that the educator contemplates and fosters collaboration and teamwork amongst the students, which mirrors the reality of the healthcare setting.

High Expectations

Clearly communicating the objectives of the simulation at a level that challenges students follows socio-cultural learning theory's premise of the existence of a ZPD (Vygotsky, 1978). As mentioned above, the ZPD is the area where the educational activity is at the appropriate level of difficulty for optimum learning to take place. The level of difficulty of the simulation should be based both on the level of the student in their educational program and individual student abilities. Ensuring that the complexity and difficulty of the simulation is appropriate may improve the student outcomes of learning, satisfaction and self-confidence.

Diverse Learning

Understanding and appreciating the diversity of student talents, learning styles and demographic characteristics is another important consideration for the educator planning a simulation activity. In many nursing education programs, there may be a mix of traditional and non-traditional students with varying levels of experience and educational preparation (AACN, 2005). Simulation activities may be flexible and accommodate heterogeneous groups of students. Best practice supports the consideration of the diversity of the student group when planning simulations.

Time on Task

Time plus energy equals learning (Chickering & Ehrmann, 1996). Good practice includes proper attention to structuring the time allotted for the simulation to allow for optimum learning and effective teaching. Clear, focused objectives with realistic time allotments and the introduction or reinforcement of only a few concepts during each session may help improve learning outcomes (Jeffries, 2006). In addition, consideration of students' knowledge and comfort level with the human patient simulator (HPS) must be taken into account when simulations are planned. It may take more time than expected to orient students to the new technologies in the simulation laboratory, and conversely students familiar with simulation may become bored with repetitive information.

Design Characteristics

Design characteristics highlighted by the NESF have been developed from the learning theory and the related literature. The NESF design characteristics are objectives, planning, fidelity, complexity/problem solving, cues/student support and debriefing. When testing the characteristics of the framework, Jeffries & Rizzolo (2006) noted improved student educational outcomes when simulations were developed taking the five design characteristics into account. The following section discusses each of these design characteristics.

Objectives and Planning

Objectives need to be very specific when planning a simulation (Medley, 2005). Raemer (2003) recommends the use of both technical and non-technical objectives. Examples of technical objectives are the development of psychomotor skills and increased knowledge. Non-technical objectives include self confidence, satisfaction and critical thinking. Jeffries (2006) cites the following planning activities as necessary to ensure the achievement of learning objectives: “identifying objectives for the experience and providing students with a time frame, guidelines for role specifications, how the simulation experience would be monitored, and how the role is related to the theoretical concepts of the course” (p. 100).

Fidelity

Simulations need to be designed to mirror the health care setting to the highest degree possible to improve student educational outcomes. Fidelity is the degree of realism or accuracy of the system used, and simulations may be categorized as low, medium, or high-fidelity (Seropian, 2004). Newer human patient simulators allow for a high degree of fidelity and as a result, simulation activities mimic reality as closely as possible. During simulation design the educator must plan to make the simulation as realistic as possible. Housing the simulator in an environment similar to the healthcare setting, having realistic equipment available, beginning the scenario with typical nurse-to-nurse shift report, and having students dress in uniform are suggestions to increase simulation fidelity (Scherer, 2003; Medley, 2005; Jeffries, 2006).

Complexity/Problem-solving

Simulations may be structured to present simple to complex nursing situations depending on the placement in the curriculum, the objectives of the simulation exercise, and student learning needs (Nehring et al., 2002). Early in the curriculum, simulations may be used to teach and practice basic psychomotor skills such as vital signs and sterile dressing changes. Complex patient scenarios with a high level of uncertainty and requiring interdisciplinary collaboration can be developed to challenge students nearing the end of their educational program.

Student Support/Cues

Cues are small pieces of information given to students at preplanned intervals in the simulation and potentially serve as a prompt or reminder for the student to act (Bremner, 2006; Jeffries, 2006). The frequency and depth of the cues are generally on a continuum with a greater number of detailed cues used with lower level students and a lesser number of cues used with advanced students. Additional cues may be available if needed, to assist students who are having difficulty completing the simulation. The appropriate level of support and cuing of information in a simulation is based on the level of student and the objectives. Appropriate student support is an integral component of the development of an educationally sound simulation exercise.

Debriefing

Critical to the design of a simulation is the planning for a period of debriefing, or reflection following the simulation (Seropian, 2003; Bremner, 2006, Jeffries, 2006). However, the amount of time and structure of the debriefing period required to achieve optimum learning outcomes has not been tested empirically. Debriefing in a non-judgmental manner allows students to reveal their patterns of thinking in a safe environment. By sharing their thinking with the educator and their peers, the group may consider other possible alternatives that may ultimately improve future clinical performance (Rudolph et al., 2006).

Outcomes

Educational outcomes identified in the NESF are developed from all three learning domains: cognitive, psychomotor, and affective (Bloom, 1957). The specific outcomes identified in the NESF are learning (knowledge), skill performance, learner satisfaction, critical-thinking, and self-confidence. Each outcome is described below.

Learning/Knowledge

Simulation is an effective strategy to teach and evaluate student learning (Bearnson & Wilker, 2005, Bremner, 2006; Jeffries, 2006). Additionally, students have reported a reinforcement of didactic knowledge through the application of prior learning in a simulation exercise coupled with an improvement on course exam scores (Comer, 2005). In contrast, Becker, Rose, Berg, Park & Shatzer (2006) found no difference in student knowledge acquisition between a group completing a simulation and control group measured with a multiple-choice exam following simulation. This finding may be related to the inability of the measure to detect improvement in the application of knowledge from the simulation exercise. Empirical evidence supporting acquisition of knowledge due to simulation is mixed, and is most likely related to a lack of reliable and valid measures.

Skill Performance

Psychomotor skill acquisition has traditionally been taught and evaluated using low to medium fidelity simulators such as static mannequins (Seropian, 2003). However, high fidelity simulation using the HPS may also be used to teach and evaluate skills in a more realistic context and allows repeated practice in a safe, supportive learning environment (Bradley, 2006). Psychomotor skill development may be measured through the use of standardized or educator developed or checklists. Skill performance practice and validation is the educational outcome with the greatest level of empirical support to date (Ravert, 2002).

Learner Satisfaction

Chickering and Gamson (1987) relate positive learner satisfaction to improved learner performance. Satisfaction with learning in simulation has been evaluated through quantitative means through the use of a 5-item Learner Satisfaction Scale ($\alpha = .94$) developed specifically for use in simulations by the NLN/Laerdal multi-site study (Jeffries & Rizzolo, 2006). A comparison of students' satisfaction ($n = 403$) between groups who used a paper and pencil case study simulation with participants in a high fidelity simulation demonstrated a significantly higher level of satisfaction ($p < .05$) among the high fidelity simulation group. Learner satisfaction as an outcome of simulation and exploration of a possible relationship between satisfaction and performance is another area for future research.

Critical-Thinking

Critical thinking is an outcome that has been measured using a variety of instruments in simulation research and findings generally support improvement in critical thinking following participation in simulation (Aronson, Rose, Anfinson & Light 1997; Jeffries, 2006; Johnson et al., 1999; Tomey, 2003; Weis & Guyton-Simmons 1998). For example, Cioffi et al (2005) found moderate support for improvement in decision-making using the think aloud technique with verbal protocol analysis. Chau et al, (2001) used the California Critical Thinking Skills Test (Facione & Facione, 1994) and found no significant improvement in critical thinking before and after participation in clinical simulation. Due to the wide variety and quality of instruments used, measuring critical thinking skill development is problematic.

Self-Confidence

It is hypothesized that practicing in the protected, safe environment of the simulation laboratory will increase students' self-confidence in the clinical setting (Bearnson & Wikler, 2005; Jeffries, 2006). Recent investigations support the development of self-confidence as an outcome of simulation (Cioffi et al., 2005; Comer, 2005). Student participants in those studies reported an increase in feelings of self-confidence to care for actual patients in the clinical setting as a result of participating in simulation.

The NESF postulates a relationship between the teacher, student and educational practices. This relationship is related to the design characteristics, the simulation intervention, and the learning outcomes. Thus, teacher factors, student factors and educational practices impact how the simulation should be designed and implemented. It is hypothesized that a simulation designed taking the aforementioned factors into account may optimize student outcomes. Further research is necessary to elucidate each of the components of the model and demonstrate how the relationships between components impacts student outcomes.

Teaching Effectiveness

As discussed above, a significant gap exists in the empirical literature describing the role of the teacher and effective teaching within the frame of clinical simulation. The role of teacher has been described throughout the anecdotal and theoretical literature as a facilitator of learning; as a helper rather than as a leader (Bain, 2004; Billings & Halstead, 2004; Hertel & Millis, 2002; Weimer, 2002). In this context, the primary role of the teacher is to design and implement creative simulations. Instructors are responsible to ensure that students have the cognitive knowledge base in addition to the affective and psychomotor abilities necessary to understand and actively participate in the simulation.

The simulation prerequisites vary depending on the objectives and educational level of the student. Simulations used for summative evaluation will have different prerequisites than simulations used to teach a basic psychomotor skill. In general, effective simulations are those which encourage students to delve more deeply into content learned in the classroom, to develop skills specific to the discipline and to apply their knowledge in a realistic context (Hertel & Millis, 2002). The role of the teacher is an integral component of effective clinical simulations. Ultimately, it is hypothesized that student participation in well designed clinical simulations will result in the transfer of cognitive, psychomotor, and affective abilities into the clinical practice setting.

Effective teaching is dependent on the coordination of several components: the objectives, the student, the content, and the teacher (McKeachie, 1997). However, McKeachie argues that “students teaching other students” is another important component which also points to the teacher serving as in a facilitative or coaching role, which fits well within the learning theories that guided the development of the NESF. Additionally, the role of the teacher in simulations may be further informed by the learning theories underpinning the NESF: constructivist, learner-centered, and socio-cultural learning theories (Bonk & Cunningham, 1998; Jeffries, 2005), and Best Practices in Undergraduate Education (Chickering & Gamson, 1987).

Boice (2004) studied a large group (n = 60-70) of highly-acclaimed educators at the undergraduate and graduate level in a variety of university settings to define best practices in college teaching. A mixed-methods approach of interview, direct observation, review of course materials, and student ratings were used to collect data. Conclusions of this study included approximately six general areas of teaching in which the participants excelled. The areas cited were: 1) “Outstanding teachers know their subjects extremely well and have at least an intuitive understanding of human learning;” 2) “Exceptional teachers treat their teaching endeavor as a serious intellectual undertaking as demanding and important as their research and scholarship;” 3) “The best teachers expect more from their students and favor objectives which embody the kind of thinking and acting expected for life;” 4) “While methods vary, the best teachers strive to create a natural, critical learning environment in which they include authentic, challenging, and important problems to challenge students within a supportive atmosphere;” 5) “Effective teachers reflect a strong trust in students, display openness with students, and above all treat students with simple decency" and 6) "All effective teachers have systematic methods to assess their teaching performance and make appropriate changes." While these findings may be useful as general guidelines for educators in clinical simulation, a lack of explanation of the research methods makes it difficult to accept these results as reliable or valid without further study.

The characteristics and behaviors of master teachers in higher education were examined by Buskist, Sikorski, Buckley & Saville (2002). The writings of master teachers about their teaching, a study of the qualities of award-winning teachers, and student perceptions of teaching were used to develop a list of the qualities of master teachers. In phase 1 of this study, undergraduates (n = 114) were asked to list at least 3 critical characteristics of a master teacher at the university level. This list contained 47 characteristics. This list was given to another group of undergraduates (n = 184) who were asked to list up to 3 specific behaviors that reflect the qualities and characteristics of master teachers on the list. Due to overlap between several characteristics, the list was collapsed to 28 qualities and behaviors for the next phase of the study.

Phase 2 consisted of providing the newly created list of 28 qualities and behaviors to a convenience sample of undergraduate psychology students (n = 916) and random sample of faculty (n = 118) from various disciplines. The participants were instructed to choose the "10 qualities/behaviors that are most important to master teaching at the college level." Data analysis revealed no differences within groups based on gender for students and faculty, or year in school for the student group. Students and faculty had 6 of the same behaviors in the top 10, but in a different order. These six qualities were: 1) realistic expectations/fairness, 2) knowledgeableness, 3) approachable/personable, 4) respectful 5) creative/interesting and 6) enthusiasm. The remaining 4 characteristics and behaviors differed between the student and faculty groups. The 4 behaviors and qualities identified by the student group related to interactions between students and faculty: 1) understanding, 2) happy/positive/humorous, 3) encouraging and 4) flexible. Faculty identified more specific classroom instructional behaviors as the remaining 4 behaviors: 1) effective communication, 2) prepared, 3) current and 4) critical thinking. However, no specific statistical information was provided regarding data analysis and as a result, it is difficult to evaluate the quality of the research.

As previously described, teaching strategies and behaviors have not been identified that reflect effective teaching within clinical simulations. An empirical literature set exists that depicts effective teaching and effective teaching behaviors in didactic and clinical contexts within higher education. However, simulations are unique and have characteristics in common with both didactic and clinical teaching. For example, simulation uses active, experiential learning that is a major component of clinical education, but simulations are conducted in a controlled setting which has similarities with classroom teaching. Theory learned in the classroom can be actively applied and synthesized in the simulation laboratory as well as in the clinical setting. Thus, clinical simulation has components in common with both clinical and didactic teaching. Since simulation is a hybrid of the two, and there is no data to define effective teaching specifically in simulation, the literature from both areas is reviewed. The majority of the literature in this review focuses on the evaluation effective teaching behaviors using student evaluations of teaching. This related literature is discussed below.

The development of an empirical knowledge base defining effective teaching behaviors and the role of the teacher in clinical simulations has been identified as a significant gap in the literature. To date, there has been no research conducted examining teaching effectiveness in this area, and subsequently no literature or measures exists to describe or evaluate the role of the teacher in clinical simulation. The NESF does postulate the role of the teacher in simulation as an integral factor of a successful simulation experience for nursing students; however evidence describing an effective teacher does not exist (Jeffries, 2005).

Effective Teaching in the Clinical Setting

There exists no definition or measure of effective teaching in simulation within the empirical literature. However, there is a substantial literature base in the clinical and classroom learning environments. Although the contexts are different, and it is recommended that measures of effective teaching be context-specific, the literature from clinical and classroom teaching add insight into teaching in simulation. Clinical teaching is related to simulations in that both are active teaching/learning strategies and simulations are designed to mirror clinical practice (Bremner, et al., 2006; Larew, Lessans, Sprunt, Foster & Covington, 2007; McCausland, Curran & Cataldi, 2004). The literature related to classroom teaching also informs clinical simulation as the learner-centered, socio-cultural and constructivist learning theories underpinning this study and the NESF support active, experiential learning in realistic contexts, which is congruent with both classroom and clinical teaching. As a result, the factors and dimensions of teaching effectiveness found in both the classroom and clinical empirical literature are reviewed and related to clinical simulation.

Clinical Teaching

Characteristics identified as effective teaching behaviors have been used to develop measures to measure effective teaching in the clinical setting (Lonser, 2006). Several instruments have been discovered in the literature in both nursing and medical education. These instruments used in nursing include: the Clinical Teacher Characteristic Instrument (CTCI) (Brown, 1981), the Nursing Clinical Teacher Effectiveness Inventory (NCTEI) (Knox & Mogan 1985, 1987), Effective Teaching Clinical Behaviors (ETCB) (Zimmerman & Westfall, 1988), the Clinical Teaching Evaluation (CTE) (Fong & McCauley, 1993), and the Teaching Effectiveness Scale (TES) (Kirschling, Fields, Imle, Mowery, Tanner, Perrin & Stewart, 1995). Measures of clinical teaching effectiveness in the medical education literature include: the Teaching Effectiveness Survey (Mourad & Redelmeier, 2006), and the Clinical Teaching Effectiveness Instrument (CTEI) (Copeland & Hewson, 2000).

The Clinical Teacher Characteristic Instrument (Brown, 1981) was developed to identify characteristics of effective clinical faculty based on faculty and student ratings. The instrument was developed based on behaviors identified through a literature review. The instrument is divided into two sections. The first section is a listing of 20 characteristics to be rated on a 5-point Likert scale ranging from “of no importance” to “of most importance”. The second section of the scale consists of participants choosing and ranking in order of importance the 5 most important characteristics of a clinical instructor (Brown, 1981). Results of the 1981 study revealed that faculty rated as the most important characteristics: professional competence, relationships with students, and personal attributes. Students ranked relationships with students first and professional competence second. No reliability information on the scale was provided. In a replication study, Bergman & Gaitskill (1990) used the CTCI to evaluate faculty and student perceptions of effective clinical instructor in a baccalaureate program. Students and faculty rated similar characteristics as most important, but in a different order.

The Effective Teaching Clinical Behaviors scale (Zimmerman & Westfall, 1988) was developed to measure the characteristics of effective clinical faculty with a sample of 281 students from three nursing programs. This scale consists 43 items rated on a 3-point Likert scale, from "not important" to "very important." Factor analysis using principle factor extraction with varimax rotation revealed a one factor solution identified as effective clinical teaching behaviors (eigenvalue = 20.61; 88.48% of explained variance). No explicit conceptual definitions or definition of effective clinical teaching behaviors were provided. Reliability was assessed using Cronbach's alpha (.97), split-half reliability (.95 to .96), and test-retest reliability [Form A (r = .94); Form B (r = .93), $p <$ .001]. To date, this instrument has not been used in any additional research.

Fong and McCauley (1993) developed the Clinical Teaching Evaluation (CTE) to assess the effectiveness of clinical teaching and to encourage nursing students to apply theory into practice based on literature review. This measure contained 30 items on a 5-point Likert scale ranging from "poor, one of the least effective teachers I know", "to excellent, one of the most effective teachers I know." Factor analysis using principle components analysis with a sample of 384 undergraduate students revealed three factors with all items loading >.50 on each respective factor: nursing competence (eigenvalue = 13.6), consideration for students (eigenvalue = 1.41), and teaching confidence (eigenvalue = 1.02). Items related to nursing competence focused on items depicting the professional nursing expertise of the instructor. Consideration for students contained items related to respecting students, instilling confidence in student and maintaining confidentiality. The teaching confidence factor related to items associated with the process of teaching in the cognitive, psychomotor and affective domains of learning.

The instrument most commonly cited and replicated in the nursing literature measuring teacher effectiveness in the clinical setting is the Nursing Clinical Teacher Effectiveness Inventory (NCTEI) developed by Mogan and Knox (1987). This instrument is designed to identify characteristics of the “best” and “worst” clinical faculty, and is based on qualitative data (Knox & Mogan, 1985) and previous research (Brown, 1981). The NCTEI is a 48 item questionnaire using a 7-point Likert scale, divided into 5 subscales: teaching ability, nursing competence, evaluation, interpersonal relationship, and personality. Summing scores from items in each subscale gives a category score. Summing all five subscale totals gives an overall score for the instructor. Higher scores imply more positive clinical teaching characteristics. Psychometric data on the scale revealed the scale to be reliable, with adequate internal consistency (α = .79-.82), test-retest scores at four weeks (r = .76-.93). The scale was reviewed by experts and was determined to have face and content validity.

Several instruments have been developed to evaluate clinical teaching effectiveness in medical education. The Clinical Teaching Effectiveness Instrument (CTEI) was developed in the U.S. by Copeland & Hewson (2000) to provide feedback to clinical faculty for self improvement and annual performance reviews. This instrument was developed based on the tailored model of clinical teaching and qualitative data. Eight concepts guided instrument development: 1) offers feedback, 2) establishes a good learning environment, 3) coaches my clinical/technical skills, 4) teaches medical knowledge, 5) stimulates independent, 6) provides autonomy, 7) organizes time for teaching and care giving and 8) adjusts teaching to diverse settings. Definitions for the concepts were not given, however examples of the concept of teaches medical knowledge (diagnostic skills, research data and practice guidelines, communication skills) and the concept of adjusts teaching to diverse settings (bedside, exam room, operating room). The instrument 15 items on a 5-point Likert scale. The authors report adequate reliability (g coefficient = .94) and validity based on expert review. The instrument was tested using ratings from medical students, residents, and fellows.

Van der Hem-Stokroos, Van der Vleuten, Daelmans, Haarman, Scherpbier (2005) replicated the reliability testing on the CTEI at a Dutch medical school, using an undergraduate medical student population as raters, with residents and staff in a surgical clerkship as educators. Results from this study supported the reliability of the CTEI as well as provided evidence for necessary sampling procedures to achieve reliable results.

The teaching effectiveness scale was developed in Canada and contains 15 items which described different characteristics of effective clinical teaching on a 5-point Likert scale (Mourad & Redelmeier, 2006). All of the scores are summed and then rescaled to achieve a rating for an individual educator ranging from 0 to 10, with a larger number indicating better teaching. Internal consistency was acceptable ($\alpha > .90$), and inter-item correlations ranged between .58-.89. Content validity was evaluated favorably with a comparison to the CTEI. Eleven of 15 items were represented in both the CTEI and the teaching effectiveness score. Mourad & Redelmeier (2006) used the teaching effectiveness score to compare educators' scores with patient care outcomes at four large

Canadian teaching hospitals. Patient records for physician participants were examined. Patients were selected who had the following diagnoses: congestive heart failure, community-acquired pneumonia, gastrointestinal bleeding, and exacerbation of chronic obstructive pulmonary disease. The authors hypothesized that there would be a positive correlation between instructor scores and patient outcomes measured by length of stay (LOS). Patients cared for by physician instructors with high teaching effectiveness ratings would have a decreased LOS. Results revealed no relationship between LOS and scores on the teaching effectiveness score, implying that educator ratings have no impact on patient LOS. However, many confounding variables, such as demographics and co-morbidities of the patient sample most likely greatly influenced the results.

Kirschling, et al., (1995) developed the Teaching Effectiveness Survey (TES), unique instrument to measure teaching effectiveness. This tool is different from other instruments available as it was designed to measure both clinical and didactic teaching effectiveness at all levels of nursing education, both graduate and undergraduates. The other measures were designed to measure clinical teaching effectiveness with under-graduate students only. The TES was developed based on a review of the literature, and rigorous two-phase psychometric testing was completed providing evidence of reliability and validity of the scale (Tables 4, 5, 6). Following psychometric testing, a 26 item survey tool with a 5-point Likert scale ranging from strongly agree to strongly disagree emerged with five subscales: knowledge and expertise, facilitative teaching methods, communication style, use of own experience, and feedback. Scores are reported for each individual item, each subscale, and total scale score. Higher scores represent more positive teaching characteristics. Despite the apparent usefulness of the TES, no additional literature could be found using the TES.

Although there are many instruments available to measure clinical teaching effectiveness in medical and nursing education, there is neither consensus on a definition of effective teaching in this setting nor consensus on the dimensions of effective teaching. Factor dimensions in the reviewed measures ranged from one to five, with the mean number of factors at four (Table 1). Copeland & Hewson (2000) identify a single dimension of effective teaching, general teaching ability. These authors argue that a one-dimensional instrument would allow easier generalizability across disciplines in a climate of increased accountability in medical education. However, a one-dimensional instrument may be more conducive to summative rather than formative evaluation which limits its usefulness. Multi-dimensional instruments are more common, but the number of dimensions varies. Through the use of multi-dimensional instruments, educators are able to assess and improve teaching performance in specific areas, and they are relevant for both formative and summative evaluations. As such, the concepts related to effective clinical teaching highlighted in this literature set are used to provide empirical data to inform the development of the SPETCS.

The NCTEI developed by Mogan and Knox (1987) was chosen from the clinical teaching evaluation instruments reviewed above for comparison with the SPETCS in order to assess criterion-related validity. The NCTEI has been used extensively in nursing clinical education in diverse settings and geographic locations. The widespread use of this instrument over a 30 year span of time made the selection of the NCTEI the logical choice for criterion-related validity assessment of the SPETCS.

Table 1

Description of Clinical Teaching Instruments

Clinical Teaching Evaluation: (Beckman & Mandrekar, 2005), Developed a 2-dimensional, 14 item 5-point Likert scale to evaluate clinical teaching. Resident physicians rated clinical faculty. Results showed 3 dimensions: interpersonal domain (3 items, loadings .74-.76); clinical teaching domain (7 items, loadings .63-.71); and efficiency domain (3 items, loadings .75 - .79). One item loaded on more than one factor and was deleted. Internal consistency reliability ($\alpha > .90$) was adequate for all domains.
Cleveland Clinic Clinical Teaching Effectiveness Instrument: (Copeland & Hewson, 2000), Is a 15 item questionnaire on 5 point Likert scale. Developed based on interviews with stakeholders (faculty, trainees, program directors, chairs), theory (tailored clinical teaching), and literature review.
Clinical Teaching Evaluation (CTE): (Fong & McCauley, 1993), A 30 item survey measured on a 5-point Likert scale. Content validated by 14 instructors (no CVI). Three-factor structure accounted for 64% total variance. Items loading > .50 on a factor were retained with 5 items deleted. Items loading on > 1 factor were placed based on opinion of researcher. Adequate internal consistency reliability ($\alpha = .97$) and test-retest reliability ($r = .85$)
Effective Teaching Clinical Behaviors (ETCB): (Zimmerman & Westfall, 1988), Scale consists 43 items rated on a 3-point Likert scale. Factor analysis using students ($n = 281$) from three nursing programs revealed a one factor solution identified as effective clinical teaching behaviors. Factor accounted for 48% of variance. Two additional factors with eigenvalues > 1 were not included because accounted for only 4% and 2.5% of variance respectively.
Ideal Nursing Teacher Questionnaire: (Leino-Kilpi, 1994), Based on Mogan & Knox (1987), and contains 52 statements measured on 5-point Likert scale. Focused on traits and required characteristics of nurse educators and is organized into 5 sections: 1) nursing competence, 2) teaching skills, 3) evaluation skills; 4) personality factors and 5) relationship with students. Face and content validity assessed by several educators. Adequate internal consistency reliability ($\alpha = .74$-$.76$) for each domain.
Nursing Clinical Teaching Effectiveness Inventory (NCTEI): (Mogan & Knox, 1987), A 48 item checklist using 5-point Likert scale, grouping teacher characteristics into 5 subscales: teaching ability, nursing competence, personality traits, interpersonal relationship and evaluation. Participants rate how the descriptive describes the teacher. Alpha ranges from .79-.92; test-retest correlation at 4-week intervals .76-.92; considered to have content and face validity per expert review.

Table 1. *continued*

Teaching Effectiveness Survey (TES):

(Mourad & Redelmeier, 2006), Scale is a 15 item survey using 5-point scale based on effective teaching behaviors identified from prior research. Used with medical student populations. Scores are summed, with highest number indicating the more effective teacher. Internal consistency is >.90 in previous studies. Content validity assessed by comparison with a Cleveland Clinic instrument. Eleven of 15 items in common.

Viverais-Dresler & Kutschke, (2001):

Researcher developed questionnaire with 3 parts: 1) demographic information; 2) 47 items rating importance with categories of professional competence, teaching ability, evaluation and interpersonal relationships. Used 7-point Likert scale. 3) same 47 items r/t satisfaction with teacher behaviors, consistency & skills. Due to length of questionnaire, four distracters were added to check if students are using same ratings consistently. Qualitative responses were obtained behaviors with highest and lowest ranking. Alpha ranged between .77-.96.

Teaching Effectiveness Evaluation Tool:

(Kirshling et al., 1995), Instrument designed to measure teaching effectiveness in undergraduate and graduate level nursing didactic and clinical courses. The instrument contains 26 items evaluating teaching effectiveness and 14 evaluating the course measured via 5-point Likert scale. Five subscales were identified: 1) knowledge and expertise, 2) facilitative teaching methods, 3) communication style, 4) use of own experiences and 5) feedback (factor loadings ranged from .62 - .92 for all subscales). Reliability of entire scale acceptable (α = .90-.94). Criterion-related validity with multiple regression demonstrated 69-74% of variance between subscales and criterion items.

Table 2

Summary of Dimensions/Factors Identified from Clinical Teaching Instruments

Author	Number of Items	Setting	Factors
Copeland & Hewson (2000); Mourad & Redelmeier (2006)	15	Medicine	General teaching ability
Dunn & Hansford (1997)	23	Nursing	Student/staff relationships Nurse manager commitment to teaching Patient relationships Facilitation of student growth
Fong & McCauley (1993)	25	Nursing	Nursing expertise Teaching competence Interpersonal skills
Mogan & Knox	48	Nursing	Teaching ability Nursing competence Personality traits Interpersonal relationships Evaluation
Lofmark & Wilkblad (2001)	Qualitative	Nursing Clinical Education	Independence Responsibility Opportunities to practice tasks Feedback
Kirshling, et al., (1995)	26	Nursing Clinical/ Classroom	Knowledge & expertise Facilitative teaching methods Communication style Use of own experiences Feedback
Viverais-Dressler & Kutschke (2001)	47	Nursing Clinical Education	Professional competence Teaching ability Evaluation Interpersonal relationships
Zimmerman & Westfall	43	Nursing Clinical	Effective clinical teaching behaviors

Classroom Teaching

Many instruments have been developed to measure teaching effectiveness in the classroom (Table 3). One of the best known is the Student's Evaluation of Educational Quality (SEEQ) developed by H. W. Marsh (1987). Marsh (1987) notes several major conclusions of the research developing and testing the SEEQ: "1) ratings consist of multiple dimensions, 2) the SEEQ is reliable and stable, 3) the results are a function of the instructor rather than of the course being taught, 4) the instrument is valid against other indicators of effective teaching, 5) results are relatively unaffected by a variety of variables hypothesized to be potential biases and 6) results are seen as useful as feedback to faculty about their teaching, by students in their course selection, and by administrators for use in personnel decisions."

The SEEQ was developed using primarily a "construct validity approach," and Marsh (1987) delineates the following assumptions that guided his research using this approach: 1) effective teaching is multidimensional, 2) there is no single criterion of effective teaching and 3) the validity and possible biases must take into account the educational context. The items for the instrument were not based on learning theory specifically, but instead were developed from a literature review, a review of existing evaluation forms, and faculty and student interviews. Next, students and faculty rated the importance of items. Faculty then judged the possible usefulness of the items for faculty feedback, and open-ended comments from students were examined for possible omissions of pertinent items from the instrument.

Table 3
Classroom Teaching Instruments

Brief Instructor Rating Scale (BIRS):
(Leamon & Fields, 2005), The BIRS is an 11 item instrument designed to evaluate undergraduate didactic teaching effectiveness. The instrument has 3 subscales characteristics of the lecturer (enthusiasm, stimulates interest, communicates clearly, knowledge of topic, rapport with students), characteristics of the lecture (organization, audiovisual/patients, syllabus/handout), and overall effectiveness (importance/worth, lecture quality, amount learned). Reliability accessed with g study across groups of students. Looked at reliability effect of crossed versus nested design with nested design requiring only 10 raters to yield a .92 g coefficient and crossed designed required 60 raters.

Heckert, Latier, Ringwald & Silvey (2006):
Developed 23 item instrument measured on 7-point scale to examine correlations between the dimensions of student ratings of teaching effectiveness and course, instructor, and student characteristics. Dimensions identified from the literature include: a global evaluation of teaching, pedagogical skill, rapport with students, difficulty appropriateness, and course value/learning. Internal consistency was adequate for all subscales (α = .70-.94). No specific mention of validity evidence. Minimal statistical analyses as correlations were used to analyze data.

Student's Evaluation of Teaching Effectiveness Rating Scale (SETERS):
(Toland & DeAyala, 2005), Is a 25 item scale using a 3 factor structure, 1) instructor delivery of course information, 2)teacher's role in facilitating instructor/student interactions, 3)instructor's role in regulating student's learning. Adequate internal consistency reliability (α = .88-.94) and criterion related validity compared to SEEQ (*r* = .13-.73) comparing each of the 3 factors with the 9 SEEQ factors.

Student Evaluation of Educational Quality (SEEQ):
(Marsh, 1983, 1987), Consists of a 35-item survey using a 5-point Likert scale. The scale has a 9 factor structure with undergraduate and graduate level students in various classroom settings and disciplines. Internal consistency reliability has been acceptable (α = .87-.98) and subscale inter-rater reliability ranged from .90-.95 for class-average responses. The factors measured by the SEEQ include: learning-value, instructor enthusiasm, organization, individual rapport, group interaction, breadth of coverage, examinations and grading, assignments and readings, and workload difficulty.

Teacher Behaviors Checklist (TBC) (Buskist, Sikorski, Buckley & Seville, 2002):
Used in populations of undergraduate students for formative and summative teaching evaluation; contains 28 items using a 5 point Likert scale Two subscales identified by factor analysis: 1) caring and supportive, and 2) professional competency and communication skills. Internal consistency reliability adequate for both subscales (α = .93-.95). Test-retest reliability from mid semester to end of the semester for the overall scale was $r = .71$, $p < .001$.

Content validity of the instrument was assumed through the item development process described above. Construct validity was assessed through factor analysis with 9 subscales emerging: 1) learning/value, 2) instructor enthusiasm, 3) organization, 4) individual rapport, 5) group interaction, 6) breadth of coverage, 7) examinations/grading, 8) assignments/readings and 9) workload/difficulty (Marsh, 1987). Criterion-related validity was measured through multi-trait multi-method analysis demonstrating acceptable correlations with other known measures of effective teaching and weak correlations with other variables known not to be related to effective teaching or potential biases, such as artificial student grade inflation. The SEEQ used class average ratings of instructor effectiveness and through development and testing of the instrument. Pooling this data without accounting for differences between the student groups is problematic as these differences are a potential confounding variable in this study.

Buskist (2002) developed the Teacher Behaviors Checklist (TBC) which matches the characteristics of effective teachers with their corresponding behaviors. The TBC operationalized personality traits into behaviors, and is based on the literature describing master teachers. The personality traits were initially identified by a group of undergraduate psychology students who listed the qualities of effective teachers. The conceptual definition of an effective teacher used in the study was one “from whom the student learned a lot and made the learning process enjoyable” (p. 188) Next, a second group of undergraduate psychology students were asked to identify specific behaviors reflective of the personality traits previously identified. A 28-item instrument was developed based on the second groups’ responses. Lastly, a third group of undergraduate psychology students and faculty rated the top 10 personality traits and behaviors from the

28 items. Findings of the comparison between students and faculty indicated a strong agreement between students and faculty ratings. The comparison of student and faculty ratings on the TBC has been replicated with similar findings in a community college context (Schaeffer, Epting, Zinn & Buskist, 2003).

Keeley, Smith & Buskist (2006) modified the previously developed TBC into an instrument to conduct formative and summative teaching evaluations. The items were amended into evaluative statements rated on a 5-point Likert scale (1 = frequent to 5 = never; 3 = no opinion). Psychology students (n = 313) at a large teaching university rated 4 different instructors. Psychometric testing of the modified instrument included, an exploratory factor analysis using principle components extraction without rotation. Three factors had eigenvalues > 1 and two factors had factor extractions greater than those found by chance using Monte Carlo simulation. As a result, two factors were retained and further factor analyzed with a one and two factor solution, using maximum likelihood extraction with nonorthogonal rotation as it was hypothesized that the 2 factors were correlated. All items loaded on one factor with loadings ranging from .47 to .79. This factor was interpreted as good teaching. A 2-factor solution was also analyzed with the criteria for item loading at .40. Twenty four items loaded on the two factors: 1) caring and supportive and 2) professional competency and communication skills. The 2-factor solution had a correlation of .73 accounting for 53% of the variance.

The 2-factor solution was retained for further study as it was argued that a 1-factor solution would be difficult to interpret. Construct validity of the instrument was assessed using a one-way ANOVA across the 4 professors using the two subscales as dependent variables. Results on both subscales were significant [F (3, 307) = 36.59, $p <$

0.001; $F(3, 308) = 19.11$, $p < 0.001$] indicating that the scale differentiated between professors. Internal consistency reliability was evident for both subscales ($\alpha = .90-.93$) and for the entire instrument ($\alpha = .95$).

Confirmatory factor analysis (CFA) was performed to test the one or two factor solutions suggested by the exploratory factor analysis (EFA) results. A one-factor, two-factor and a hybrid model all postulated that the two subscales may be components of a higher order factor which may capture the items that did not load originally in the EFA procedures. Undergraduate psychology students from 5 different course sections with 5 different faculty completed the instrument at two time points, mid-semester ($n = 313$) and the final week of the semester ($n = 322$). The 5-point Likert scale was amended, with 1 = always exhibits behaviors, 2 = frequently, 3 = sometimes, 4 = rarely, 5 = never.

Model fit was assessed using three indices: the normed fit index (χ^2), comparative fit index, the root mean square error of approximation (RMSEA). The authors argue that using multiple indices of fit rather than a single index provides more complete evidence of construct validity. Results supported all three solutions, with the χ^2 statistic having the lowest value on the two-factor solution (one-factor model $\chi^2 = 1340.474$, $df = 350$; two-factor model $\chi^2 = 913.213$, $df = 251$; hybrid model $\chi^2 = 1169.269$, $df = 348$). Results of model fit testing were similar at both time 1 and time 2. However, it is difficult to interpret the differences in the χ^2 because the models are not nested and the differences could be due to chance.

Test-retest reliability was assessed using Pearson's correlation coefficient. In addition, a regression analysis of the difference between mid and end of semester scores was analyzed to account for the magnitude of change that was hypothesized to occur over time. Results provide evidence suggestive of adequate test-retest reliability with Pearson's $r = 0.71$ ($p < 0.001$) for the total scale, $r = .68$ ($p < 0.001$) for the caring and supportive subscale, and $r = 0.72$ ($p < 0.001$) for the professional competency and communication subscale. Regression analysis was performed using transformed deviation scores from mid-semester (score minus its mean) with the intercept becoming the mean of the end of semester scores and the slope determined the direction and magnitude of change from mid-semester. A slope of 1 corresponded with an increase of 1 point on the Likert scale. All individual items had positive slopes ranging from .02-.05 ($p < .001$). The slope of the total scale = .65 ($p < 0.001$), the caring and supportive subscale = .57 ($p < 0.001$), and the professional competency and communication subscale = 0.71 ($p < 0.001$). The results of the regression analysis indicated that scores improved about one-half to one point by the end of the semester, and that the professional competency and communication scores improved to a greater degree than the caring and supportive subscale scores.

Results of the psychometric testing provided initial support of the internal consistency and test- retest reliability of the revised TBC as an evaluative tool. Additionally, the hypothesis that the subscale scores would increase from mid to the end of the semester was supported. The results of the goodness of fit CFA did not provide support for one model over another and the authors suggested using all three models in future research using the TBC. One significant limitation evident was the nested design of the study as each section of students rated only 1 instructor rather than rating all instructors. Results of each section of students should have been analyzed separately in addition to the pooled results reported in the study. External validity was an issue as it would not be possible to generalize to other student types and other educational settings.

Although measures of clinical and classroom teaching were developed for use in those respective contexts, the dimensions or factors of effective teaching identified in this body of research (Table 4) and the common pitfalls of measurement that occurred in many of the reviewed studies served as a guide to inform design of this research and development of the SPETCS. Additionally, the use or modification of an existing instrument would significantly threaten the internal validity the study as the construct of teaching effectiveness differs based on the learning environment. The dimensions identified for use in the SPETCS were developed based on those from the literature that were congruent with simulation and the learning theories underpinning this research.

Table 4.
Summary of Dimensions/Factors Identified from Classroom Teaching Instruments

Author	Number of Items	Setting	Factors
Marsh (1987)	35	Undergraduate & graduate	Learning/Value Instructor enthusiasm Organization Individual rapport Group interaction Breadth of coverage Examination/grading Assignments/reading Workload/difficulty
Buskist et al. (2002)	28	Undergraduate	Caring/Supportive Professional competency/ communication skills
Heckert, Latier, Ringwald, Silvey (2006)	4 dimensions/ global rating	Undergraduate	Global teaching effectiveness Pedagogical skill Rapport with students Difficulty appropriateness Course value/learning
Leamon & Fields (2005)	11	Undergraduate Medical	General instructional skill Lecturer Lecture/overall
Witcher et al.(2003)	Qualitative	Undergraduate	Dedicated Fair/Competent Knowledgable Enthusiastic
Toland & DeAyala (2005)	24	Undergraduate	Instructor delivery of course information Teacher's role in Facilitating instructor/ student interactions Instructor's role in regulating student's learning

Synthesis of Effective Teaching Literature in Clinical and Classroom Settings

Several common issues were prevalent throughout the literature set. Most notable and problematic was that most of the studies were atheoretical and lacked conceptual definitions of effective teaching and its related dimensions or constructs. The lack of definitions and theoretical frameworks makes interpretation, synthesis and replication of these studies difficult. Several studies used a literature review to create their measure, but no learning theories were explicitly cited, with the exception of tailored teaching (Copeland & Hewson, 2000). Empirical data generated without a precise and detailed definition of the target construct calls into question the ability of the research to contribute to the literature (Clark & Watson, 1995).

Evidence of reliability reported for the instruments varied. The majority of instruments analyzed internal consistency reliability using Cronbach's alpha, which is the most generally accepted analysis (DeVellis, 2003). Two of the instruments in medical education used generalizability studies to assess reliability (Copeland & Hewson, 2000; Leamon & Fields, 2005). Test-retest reliability was tested in relatively few studies, with many using Pearson's correlation coefficient. The intra-class correlation coefficient has been highlighted in the recent literature as a better method of analysis than Pearson due to the fact that a test-retest examines correlation between the same variables (instruments) and Pearson was developed to examine correlation between different variables (DeVon 2007; Yen & Lo, 2003).

Issues with validity were also pervasive through the literature set. Content validity was generally claimed in all of the studies, but the methods used to assess content validity were rarely specified. Those that did specify used an expert panel to review items, however the composition and selection criteria for panel members were generally missing. Details of how items were analyzed in the instrument development studies were absent as well. Dimensionality of instruments and construct validity were generally evaluated through the use of either exploratory or confirmatory factor analysis. Exploratory factor analysis instruments were assessed using principle components factoring. Newer recommendations are to use principle axis factoring when completing an EFA (Netemeyer, et al., 2003).

Pooling of data from different courses, teachers, and level of student commonly occurred throughout the literature without controlling for this potentially confounding variable. Two studies addressed this issue. Toland (2005) used multilevel factor analysis to control for the effect of using different student groups. Beckman & Mandrekar (2005) used a homogeneous sample of medical residents to diminish the effects of pooling student groups. However, the majority of studies did not account for these differences which may have a significant impact on the results by increasing the possibility of Type I error (DeVon, 2007).

Convergence or criterion-related validity was sparsely reported in the literature. Toland (2005) compared results on the SETERS subscales with the subscales on the SEEQ using Pearson's correlation coefficient (r = .13-.73). Results from the same sample on related instruments should be significantly correlated with one another to demonstrate evidence of criterion-related validity. Other studies used criterion items, which were statements to assess teaching effectiveness globally, or based on instrument subscales. Kirshling et al., (1995) compared three criterion items with results on each of five subscales of the instrument using multiple regression analysis. The summed subscales accounted for 69-74% of the variance compared to the criterion items.

In sum, several instruments developed to evaluate teaching effectiveness in the classroom and clinical settings were reviewed. Only a single instrument was designed to assess teaching in both settings (Kirshling, et al., 1995), and no instrument was found to measure effective teaching in simulated learning environments. A number of authors specifically argued for the creation and use of teaching effectiveness measures that are context specific (d'Apollonia & Abrami, 1997; Beckman, Ghosh, Cook, Erwin & Mandrekar, 2004; Beitz & Weiland, 2005). Reliability and validity evidence was generally assessed through the use of Cronbach's alpha and factor analysis. Evidence of test-retest reliability, criterion-related validity, and content validity were reported inconsistently. The next section delineates the mechanisms to address the issues of measuring effective teaching in the proposed study.

First, a conceptual definition of effective teaching specific to simulated learning environments has been developed. Findings from the literature review, socio-cultural, constructivist, and learner-centered learning theories and the NESF guided the development of the SPETCS. The relationship between the theory, definition, and the SPETCS has been explicitly stated. The SPETCS was designed specifically for use in simulation contexts and was multidimensional to allow both formative and summative feedback to faculty. The sampling, reliability and validity issues discussed above are addressed in the psychometric design of this study and are discussed in detail Chapter 3.

3. METHODOLOGY

This chapter describes the methods and procedures that were used to develop and to test the Student Perception of Effective Teaching in Clinical Simulation (SPETCS). Details of the research design, sample selection criteria, ethical considerations, instrument development, variables measured, and data analysis methods are discussed below.

Design

A quantitative, cross-sectional, descriptive design was used in this study which consisted of two components: 1) instrument development and 2) assessment of the psychometric properties of the newly developed instrument. A descriptive design was chosen based upon the ability of descriptive research to describe a particular phenomenon in a specific population in an area about which little is known (Kerlinger & Lee, 2000). The Nursing Education Simulation Framework (Jeffries, 2005), a conceptual framework developed based upon learner-centered, socio-cultural, and constructivist learning theory provided the theoretical basis for this study (Bonk & Cunningham, 1998). The latent variable of interest in this study was teaching effectiveness within the context of clinical simulation. The constructs describing the domain of teaching effectiveness have been developed through a literature review and were content validated by nurse experts in clinical simulation as designated by the National League for Nursing.

The use of high-fidelity clinical simulations in nursing education is a relatively new teaching-learning strategy. As discussed in the literature review, there is a significant gap in the literature examining effective teaching in clinical simulations. In addition, no instrument exists to measure teaching effectiveness within the clinical simulation environment. As a result, an instrument has been developed and tested to measure effective teaching behaviors in clinical simulation contexts. The SPETCS can serve as a tool for assessment, evaluation, and feedback in the ongoing professional development of nursing educators who use clinical simulations in the teaching learning process.

Phase 1 - Instrument Development

Instrument development is a two-stage process (Lynn, 1986). The first is the developmental stage and the second is the judgment-quantification stage. In the developmental stage, the domain of interest is identified beginning with the creation of a clear conceptual definition and a review of the literature. The conceptual definition of effective teaching in clinical simulation for this study is the degree to which the teaching strategies and behavioral characteristics of the instructor promote student achievement of the learning outcomes specified in the simulation experience. The conceptual definition has been reviewed by four doctoral prepared nurses, three with expertise in nursing education, and one with expertise in research methods.

Next, items for the SPETCS were developed based on the literature and conceptual definition of the dependent variable. Constructs identified in the literature review and congruent with the conceptual definition were used as the basis to operationalize effective teaching strategies and behaviors based on the conceptual definition. Items were clustered around one of these constructs. The constructs serving as the basis for the measure were:

1. Facilitator-Learner Centered: The role of the teacher is to provide the structure for and assist, or coach student learning. The focus is squarely on the learner in an environment conducive to learning. Instruction is designed to allow

students to control and be responsible for their learning (Vygotsky, 1978; Weimer, 2002).

2. Feedback and Debriefing: Feedback is the constructive and non-judgmental appraisal of performance; it may be formative or summative. The type and amount of feedback is assessed (Bienstock, et al., 2007; Mogan & Knox, 1987).
3. Teaching Ability: Ability to design, to implement, and to evaluate simulations effectively. Specifies a knowledge base and comfort level with simulation technology (Jeffries, 2007).
4. Modeling: Modeling is the process of offering behavior for imitation. Modeled activities can be transformed into images and verbal symbols that guide subsequent performance. Faculty or more advanced peers may serve as a model (Dunphy & Dunphy, 2003).
5. Interpersonal Relationships: Faculty to student and student to student, and group interactions/communication. These interactions are based on mutual respect, fairness, and confidentiality (Copeland & Hewson, 2000; Mogan & Knox, 1987).
6. Expectations: The degree of difficulty of the simulation is appropriate for optimal student learning to occur. Generally is based on the level of students and assessment of students' zone of proximal development (Sanders & Welk, 2005; Vygotsky, 1978).
7. Organization: The simulation experience (simulation + debriefing) flowed smoothly. The simulation design, implementation, and evaluation were

orderly, functional, and well structured into a coherent whole. http://dictionary.reference.com/browse/organized.

8. Enthusiasm: The instructor is dynamic and energetic. Demonstrates excitement and interest towards subject (Marsh, 1987).
9. Cuing: Verbal, non-verbal, or written information or data provided to the learner to spur their thinking in order to make meaning of the situation. (Alfaro-LaFevre, 2004; Larew, et al., 2007; Pesut & Herman, 1999).
10. Questioning: Questioning calls for an active linguistic and cognitive response, and it provokes creations by the learner (Dunphy & Dunphy, 2003). There are two kinds of questions, those that assess and those that assist. Appropriate questioning allows the educator to know how much assistance is needed to optimize learning (Sanders & Welk, 2005).

Items were structured based upon the following recommendations of Baroque and Fielder (1995): 1) items were written as complete sentences, 2) abbreviations, jargon, and technical expressions were avoided and 3) loaded or biased phrases were not used. Demographic data to be obtained from participants was age, gender, ethnicity, race, current grade point average, other college degrees, previous experience with clinical simulations and years of experience working in health care.

The judgment-quantification stage began with content validation of the instrument. One of the primary objectives of this stage was the quantification of content validity of individual items through the calculation of the content validity index (CVI), as recommended by Lynn (1986). The CVI is a quantitative method that provides evidence of content validity of the instrument which is the “determination of the content representativeness or content relevance of the items in an instrument" (Lynn, 1986). The CVI was calculated based on expert panel ratings of each individual item and on the entire scale. The expert panel for this study consisted of geographically diverse nurse experts (n = 7) in clinical simulation as designated by the National League of Nursing. The expert panel rated each item on a Likert scale according to Lynn's criteria, for relevance to teaching effectiveness. Consideration for item revisions, additions, or deletions was based on panel recommendations.

The expert panel was provided with the conceptual definition, the theoretical basis, and the measurement model for this study (Grant & Davis, 1997). The panel of simulation experts (n = 7) rated the content relevance of each item using a 4-point Likert scale: 1) not relevant (content invalid), 2) somewhat relevant (slightly invalid), 3) quite relevant (slightly valid) and 4) very relevant (content valid). Next, the item scores were collapsed into dichotomous categories; scores of 1 or 2 indicated an invalid item, and scores of 3 or 4 indicated a valid item.

Once the categories were collapsed, the CVI for each item and the entire instrument was determined. The CVI for an individual item was calculated as the proportion of valid items to the number of experts. The CVI for the entire scale was measured through the calculation of the proportion of items receiving a score of 3 or 4 to the total number of items in the scale. Based on criteria supported by Lynn (1986) and Wynd (2003), items require a CVI of ≥ .78 to be considered content valid. For the entire scale, the criterion for an adequate CVI was: 1) .70 = average and minimal value necessary to determine agreement, 2) .80 = adequate agreement and 3) .90 = good agreement.

Phase 2 – Psychometric Testing of Instrument

Following development of the instrument and establishment of content validity, two nursing students examined the study materials to determine readability. Next, a pilot study to determine the feasibility of the research protocols was conducted. A convenience sample of 29 students participated in the identical clinical simulation with the same instructors as was planned for the larger study. All participants provided demographic and situational information and completed the three study instruments. As a result of the pilot study, the research plan for the larger study was validated with only minor changes needed to the format of one of the instruments.

After the pilot study, the SPETCS underwent rigorous reliability and validity testing with a large sample (n = 121) of undergraduate students participating in a high-fidelity clinical simulation. The SPETCS was tested for evidence of internal consistency reliability and test-retest reliability. Individual items were analyzed using Ferketich's (1991) criteria. Ferketich's criteria have been explained in the data analysis section of this chapter. Dimensionality of the instrument was tested through the use of factor analysis using principal axis factoring (Netemeyer, et al., 2003). Criterion-related validity was assessed by calculating a correlation between the SPETCS and participant scores on two well-known student evaluation of teaching effectiveness instruments, the Student Evaluation of Educational Quality (Marsh, 1987) and the Nursing Clinical Teaching Effectiveness Inventory (Mogan & Knox, 1987). The procedures for data collection and analysis are presented in detail below.

Sample

A purposive sample of faculty experts (n = 9) in clinical simulation designated by the National League for Nursing were asked to complete the content validation component of the study. Seven experts completed the content validation survey. Lynn (1986) recommends a sample size of at least five experts to decrease the probability of a chance agreement during content validity assessment. However, DeVon, et al. (2007) notes that complete agreement on items must exist if the sample size has 7 or fewer experts assessing content validity. Thus, the planned sample size of nine experts was appropriate, and with this sample size, seven out of nine panelists must be in agreement to achieve an acceptable CVI. Two nursing students reviewed the instrument for clarity and readability.

A convenience sample of undergraduate nursing students (n = 29) participated in the pilot testing of the instrument and research protocols. A larger convenience sample of undergraduate students enrolled in the same nursing course (n = 121), who participated in a high-fidelity clinical simulation, were used for initial psychometric testing of the newly-developed instrument. The students were taught by one of two different instructors. The sample size was determined based upon recommendations in the literature regarding an adequate sample size for exploratory factor analysis. Exploratory factor analysis with principal axis factoring (Netemeyer, et al., 2003) was used to provide insight into the dimensionality of the SPETCS. However, opinions differ regarding an appropriate sample size for factor analysis, and in general large samples are preferable as the results tend to be more stable (DeVellis, 2003; MacCallum, Widaman, Zhang & Hong, 1999; Munro, 2005; Tabachnik & Fidell, 2001). A common rule of thumb cited in the literature recommended a sample size of at least 5 participants per variable or 50 participants greater than the total number of variables (DeVon, 2007). Thus, the planned sample size of 100 was both pragmatic and adequate for initial psychometric testing of a newly developed instrument. In addition, the entire sample was asked to complete the SPETCS instrument 2 weeks following initial testing to provide evidence to support test-retest reliability. The procedures for both the item development and psychometric testing phases of this study are described below.

Procedure

Expert panelists for content validation were contacted by electronic mail to request participation. A cover letter was sent describing the aims, theoretical underpinnings, the conceptual definition of effective teaching, and the measurement framework of this study. Instructions for completion and return of the survey instrument were included with the content validation materials (Appendices A & B). Panelists were asked to complete the instrument and submit it electronically within two to three weeks. The responses of the panel were used to quantitatively assess the relevance of items to teaching effectiveness and to make necessary revisions to the SPETCS.

Following approval from the Indiana University Purdue University Institutional Review Board, undergraduate nursing students who were identified as participants in clinical simulations as a component of their medical-surgical clinical nursing course were recruited by the researcher. Students typically rotated through simulations to enhance hospital-based clinical experiences at specific time points during the semester. Two instructors facilitated the high-fidelity simulation that formed form the basis for this study. The researcher approached students as a group one week prior to the simulation experience regarding participation in the study. The purpose and the voluntary nature of the study were explained to potential participants in detail, and the informed consent form was signed.

Immediately following the debriefing component of the simulation experience, participants completed the three survey instruments: 1) the Student Perception of Effective Teaching in Clinical Simulation, 2) the Students' Evaluation of Educational Quality (Marsh 1987) and 3) the Nursing Clinical Teaching Effectiveness Inventory (Mogan & Knox, 1987). All materials were provided to participants in a paper and pencil format. Two weeks following the initial data collection, participants were asked to reflect back on their simulation experience from two weeks prior and complete the SPETCS a second time for analysis of test-retest reliability.

Human Subjects Approval

Approval for the study was obtained from the Indiana University Purdue University Institutional Review Board (IUPUI-IRB). The purpose, risks, and benefits of this study were explained to potential student participants by the researcher. Students were informed of the voluntary nature of the study, that participation may be withdrawn at any time, and that participation in this study had no impact on their grades for the course. Participants signed an informed consent form one week prior to completion of the self-report survey instruments. All surveys were completed anonymously in a paper and pencil format. As there was no intervention involved in this study, application was made to the IUPUI-IRB an Exempt study.

Confidentiality and anonymity was protected through the use of a numerical coding system on all instruments. The master list of all study participants and research data was managed by the principal investigator and student researcher only. This information was located on the hard drive of the student's computer. The computer used to store research data was password and firewall protected. The PI and student researcher, had access to coded data for statistical analysis.

Variables and Instruments

Student Perception of Effective Teaching in Clinical Simulation

The Student Perception of Effective Teaching in Clinical Simulation (SPETCS) was developed and psychometrically tested in this study (Appendix A). The scale had two response scales, Extent and Importance. Each response scale used a 5 point Likert scale. The Extent response scale asked students to rate their extent of agreement with each survey item related to particular teaching behaviors or strategies used during the simulation: 1) strongly disagree, 2) disagree, 3) neutral (neither agree or disagree), 4) agree and 5) strongly agree. The Importance response scale was scored based on student perceptions of how important the teaching behavior or strategy used during the simulation was towards meeting the simulation learning outcomes: 1) not important, 2) slightly important, 3) moderately important, 4) very important and 5) extremely important. Individual item scores and summed scores were analyzed. Individual items with the highest mean scores in the Extent response scale were those behaviors or strategies most frequently used in the simulation. Item means in the Importance response scale represented student participants' perceptions of which instructor behaviors or strategies were most important to meeting the specific simulation objectives.

Additionally, summed subscale scores were examined with higher total scores on the Extent response scale indicative of more frequent use of teaching strategies and behaviors. Summed scores on the Importance response scale were indicative of the importance of all of the teacher behaviors and strategies used towards meeting simulation objectives.

Nursing Clinical Teaching Effectiveness Inventory

The NCTEI (Appendix C) developed by Morgan & Knox (1987) is the most widely used instrument measuring nursing clinical teaching effectiveness. The scale consists of a 48 item checklist using 5-point Likert scale grouping teacher characteristics into 5 subscales: 1) teaching ability, 2) nursing competence, 3) personality traits, 4) interpersonal relationship and 5) evaluation. Summing scores from items in each subscale gives a category score. Summing all five subscale totals gives an overall teaching effectiveness score for the instructor, with higher scores indicative of more positive teaching characteristics. The instrument was developed from prior qualitative data describing student perceptions of the ‘best’ and ‘worst’ clinical instructors (Mogan and Knox 1983) and previous research (Brown, 1981). Psychometric analysis of the scale indicates acceptable internal consistency reliability (α = .79-.92), test- retest correlation at 4 week intervals (r = .76-.92), and is considered to have content and face validity (Mogan & Knox, 1987).

Student’s Evaluation of Educational Quality

The SEEQ (Appendix D) is a well-known, widely used instrument to measure students’ evaluation of classroom teaching and consists of a 35-item survey using a 5 point Likert scale (Marsh, 1983, 1987). The scale has a 9 factor structure that has been supported by many studies with undergraduate and graduate level university students in various settings and disciplines (Marsh 1987; Marsh & Hocevar, 1991). Internal consistency reliability has been acceptable (α = .87 .98), and subscale inter-rater reliability ranged from .90-.95 for class-average responses.

The factors measured by the SEEQ include:

1. Learning-value (4 items)
2. Instructor enthusiasm (4 items)
3. Organization (4 items)
4. Individual rapport (4 items)
5. Group interaction (4 items)
6. Breadth of coverage (3 items)
7. Examinations and grading (4 items)
8. Assignments and readings (4 items)
9. Workload difficulty (4 items)

Data Analysis

Data Screening Procedures

Data cleaning began by ensuring the accuracy of data entry into the SPSS statistical software package (SPSS, Chicago, IL). All data was double checked for accuracy prior to statistical analysis. Additionally, data was screened and corrected using SPSS statistical software to assess for normality, linearity, homoscedasticity, skewness, kurtosis and outliers (Tabachnik & Fidell, 2001). Missing data and the patterns of the missing data were analyzed. Correcting missing data through the input of items was thoughtfully considered as there were several possible actions that may be taken to rectify missing data, each with its own implications. Additionally, multicolliniarity and singularity of variables was assessed and no corrective steps needed to be taken (Tabachnik & Fidell (2001).

Specific Aims and Hypotheses

Data was analyzed for each specific aim and hypothesis using SPSS statistical software (SPSS, Chicago, IL). The level of significance for each analysis was $p < .05$. The aims and hypotheses for this study and the statistical analysis plan for each are discussed below.

Specific Aim 1. To develop the Student's Perception of Effective Teaching in Clinical Simulation (SPETCS) scale and determine the degree of relevance of individual items and the overall scale using Lynn's (1986) content validity index (CVI).

Hypothesis 1a. The SPETCS scale items demonstrate a CVI of ≥ .78. Using a panel of nine experts, seven need to be in agreement regarding the relevance of each item to obtain an acceptable CVI of ≥ .78 (Lynn, 1986).

Each item was analyzed and those demonstrating a CVI of .78 or above were retained in the scale. Items with scores below the set threshold were evaluated for possible deletion or revision.

Hypothesis 1b. The SPETCS total scale demonstrates a CVI of ≥ .78, which is the proportion of total items judged content valid.

A content validity index was calculated on the total scale, which is the proportion of items in the scale is deemed content valid to the total number of items in the scale. Again, a CVI of .78 is the threshold recommended by Lynn (1986). Following item revisions, possible additions and/or deletions as recommended by the expert panel, the total scale CVI should approach 1.0.

Specific Aim 2. Evaluate the psychometric properties of the SPETCS scale.

Hypothesis 2a. The SPETCS scale items demonstrate means near the center of the scale (3), the range of standard deviations indicates variability in the data, and floor and ceiling effects are less than 10% (DeVellis, 2003).

Descriptive statistics with a focus on the mean individual item scores were analyzed. Emphasis was placed on whether the mean items scores were near the midpoint of the scale and analysis of the variability in the scores compared to the mean as reflected by standard deviations. Floor and ceiling effects of the data were examined and data was assessed for outliers.

Hypothesis 2b. Item-to-total correlations for items in the SPETCS scale are ≥ .30 among student participants in simulation (Ferketich, 1991).

Ferketich (1991) recommends that scale items correlate with one another at a level ≥ .30 and < .70. Items measuring the same construct should correlate with one another. If an item score is < .30, that item may not be measuring the construct of interest. In contrast, if an item scores > .70 it may be redundant. Items with average inter-item correlations outside these parameters were evaluated for deletion from the scale. Corrected item-to-total correlations were also analyzed. It was hypothesized that the greater the correlation between the item and the total, the better the item. Nunnally (1978) recommends correlations above .30. If an adequate number of items meet these guidelines, no changes were to be made to the scale. Analysis of the internal consistency of the scale is described below.

Hypothesis 2c. The SPETCS scale demonstrates evidence of internal consistency reliability with Cronbach's alpha values $\geq .70$ among student participants in simulation (Carmines & Zeller 1979; DeVellis, 2003; Nunnally, 1978)

Internal consistency relates to the homogeneity of items in a scale (DeVellis, 2003). Based on classical measurement theory, scales are developed to measure a single construct and as such, all of the scale items should have a strong relationship to one another. A scale is internally consistent to the degree that the items in the scale are highly inter-correlated (DeVellis, 2003). The most widely used statistic to examine the internal consistency of an instrument is Cronbach's alpha (α). The alpha statistic specifies the portion of total variance from the item set that is unique, and subtracts this from 1 to determine the amount of variance that is communal. Next, this value is multiplied by a correction factor to account for the number of items in the scale (DeVellis, 2003). The SPETCS demonstrates evidence of internal consistency reliability of the total scale and possible individual dimensions (subscales) through an acceptable Cronbach's $\alpha > .70$ (Carmines & Zeller, 1979; Nunnally, 1978).

Hypothesis 2d. Evidence of temporal stability of the SPETCS scale is provided by 2 week test-retest reliability with an intra-class correlation coefficient $> .60$ among student participants in simulation (Shrout & Fleiss, 1979; Yen & Lo, 2002).

Stability of the SPETCS was assessed through calculation of the intra-class correlation (ICC) of participant's scores taken two weeks apart. The entire sample of students was asked to complete the SPETCS at time 1 and time 2. To complete the questionnaire at time 2, participants reflected back on their simulation experience that occurred 2 weeks prior.

Yen & Lo (2002) argue that the ICC is more appropriate to assess test-retest reliability than the Pearson's product-moment correlation. Pearson's *r* is designed to test the correlation between 2 variables, and as a consequence it is theoretically invalid to use this statistic to assess correlation between items measuring the same variable. In addition, the ICC reflects the presence and magnitude of systematic error while Pearson's *r* does not. The approach outlined by Yen & Lo (2002) was used to assess the ICC for test-retest reliability. First, teaching effectiveness was tested twice at two week intervals. An ICC of 0-.02 indicates slight agreement, .21-.4 indicates fair agreement, .41-.60 indicates moderate agreement, .61-.81 indicates substantial agreement, and .81-1.0 indicates near perfect agreement (Landis & Koch, 1977). An ICC > .60 is considered to be appropriate for this study as this value is indicative of substantial agreement between scores on the first and second administrations of the instrument.

> *Hypothesis 2e*. The dimensionality and initial evidence of construct validity of the SPETCS with factor loadings of .32 (Tabachnik & Fidell, 2001), and above for each domain is determined through exploratory factor analysis using principle axis factoring among student participants in simulation (Netemeyer, Bearden & Sharma, 2003).

The purpose of the exploratory factor analysis (EFA) is to identify the underlying dimensions in the newly created instrument. These dimensions or factors are subsets of the entire pool of items. The items in each subset are correlated with one another, but are relatively independent of items in other subsets (Tabachnik & Fidell, 2001). These factors reflect the number of variables underlying the set of items, which condenses information. As a result, variation in the scale can be accounted for using a smaller set of variables (DeVellis, 2003). Additionally, EFA may assist with defining each of the identified factors accounting for the variation among the larger item set (Tabachnik & Fidell, 2001). EFA examines the shared variance between items in the scale as it is hypothesized that if the underlying latent constructs, or factors explain the observed variables, it is important to focus on the variance shared among the observed variables (Tabachnik & Fidell, 2001)

Initially, an independent samples Student's *t*-test was calculated to analyze the mean responses from groups taught by the two different instructors. It was hypothesized that there would be no significant difference ($p > .05$) of mean scores between the two groups. These groups were in the same level in the undergraduate program and an identical simulation scenario was used for all students. Data from both instructors was to be combined for EFA if the *t*-test was not significant.

The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) was assessed to evaluate the appropriateness of performing factor analysis (Munro, 2005). Tabachnik & Fidell (2001) recommend a level of ≥ .60 is required for "good FA." If the KMO requirement is met, an initial EFA will be run. The first step to interpret the results is the examination of eigenvalues. An eigenvalue is the amount of variance of the original variables explained by the factors, or the amount of information captured by a factor (DeVellis, 2003; Tabachnik & Fidell, 2001). On the initial EFA, Kaiser's rule will be used and factors with eigenvalues > 1 will be extracted. DeVellis (2003) and Munro (2005), argue against strict adherence to Kaiser's rule because using eigenvalues > 1 may allow the extraction of too many factors. Both recommend using a scree plot to make decisions regarding the determination of the number of important factors. The scree plot is a graphic representation of factors plotted against the eigenvalues and the plot will be examined to assess where a line drawn through the points changes slope (Tabachnik & Fidell, 2001). Factors found above the point where the slope changes are factors to be extracted. Conversely, any factors below this point are those with low eigenvalues that do not contribute significantly to the shared variance and is part of the scree along the bottom of the graph. For this study, both eigenvalues and the scree plot will be used to determine the number of factors that are important to extract.

Following initial analysis, the data was rotated using orthogonal varimax rotation to improve the interpretability of the results. Orthogonal rotation assumes that factors are uncorrelated with one another, and varimax is the most widely used rotation method (DeVellis, 2003; Munro, 2005). Varimax rotation maximizes the variance so that a variable in the data set is either highly or lowly correlated with each factor. The goal of rotation is to have each variable load highly on one factor and lowly on the others, which is known as achieving simple structure (Tabachnik & Fidell, 2001). Individual factor loadings will be analyzed, with loadings ≥ .32 being interpreted. Factor loadings provide evidence related to the degree to which the variable is a pure measure of a particular factor (Tabachnik & Fidell, 2001).

Exploratory factor analysis was used to assess the dimensionality of the SPETCS. The factors or dimensions underlying the construct of effective teaching in simulation were identified. Additional subscales were identified at this point. EFA was also used to revise the instrument based on the outcome of the initial EFA with varimax rotation.

Following instrument revision, EFA procedures and the reliability procedures discussed were repeated to ensure the scale remains reliable.

> *Hypothesis 2f.* The SPETCS scale demonstrates evidence of criterion-related validity as evidenced by significant ($p < .05$) correlations with the Student Evaluation of Educational Quality (SEEQ, Marsh, 1987) and the Nursing Clinical Teacher Effectiveness Instrument (NCTEI; Mogan & Knox, 1987).

Criterion-related validity relates to an empirical association or correspondence with a criterion or gold standard (Carmines & Zeller, 1979). In contrast to other types of validity assessment, criterion-related validity requires no theoretical understanding of the correlation between the instrument and the criterion (DeVellis, 2003). A measure is criterion-valid when the data from the instrument and another variable are correlated and the “strength of the correlation substantially supports the extent to which the instrument estimates performance” (DeVon, et al., 2007).

Problematic for this study and a primary rationale of the need to develop the SPETCS is the lack of a gold standard instrument to use to establish criterion related validity in simulation environments. However, two widely used instruments measuring effective teaching in the college classroom and the clinical areas were used to demonstrate criterion-related validity. The SEEQ (Marsh, 1987) and the NCETI (Mogan & Knox, 1987) were completed by all participants during the initial administration of survey instruments immediately following the clinical simulation experience. Individual mean scores of the SPETCS and the SEEQ were compared and a separate comparison was made between the NCTEI and the SPETCS scores. Confidence intervals were also calculated due to the fact that a moderate correlation may provide strong validity support if the *r* value falls within a narrow (95%) confidence interval (DeVon, et al., 2007). Significance of the correlations were set at $p < .05$. Detail regarding the instruments is provided in the variables and instruments section of this chapter.

Specific Aim 3. Determine which teaching strategies/behaviors are most frequently used in clinical simulations based on perceptions of student participants.

Mean scores in the Extent response scale were calculated and analyzed for each item. Items were rank ordered from the highest to lowest mean. Teaching strategies/behaviors with the highest means were identified as those most frequently used in the simulation as perceived by student participants.

Specific Aim 4. Determine which teaching strategies/behaviors are most important to facilitate achievement of specified simulation outcomes based on ratings of student participants in simulation.

Mean scores of items in the Importance response scale were calculated and analyzed. Items were rank ordered from highest to lowest based on mean scores. Teaching behaviors and strategies with the highest means were identified as those identified as most important to the achievement of simulation objectives as perceived by student participants.

Specific Aim 5. Determine the relationship between student participant's demographic and situational variables and their responses on the Extent and Importance response scales of the SPETCS. Determine if scores between the two student groups who had different instructors are similar.

The relationship between participant demographic and situational variables and the responses on the response scales of the SPETCS were calculated using a multiple regression model. Two regression equations were analyzed, one with the Extent response scale and one with the Importance response scale as dependent variables. Demographic and situational variables were entered as the independent variables. Continuous variables were screened prior to inclusion into the regression equation through the analysis of Pearson product-moment correlation coefficient with the two response scales as dependent variables. Discrete variables were screened by analysis of the MANOVA univariate F values using the response scales as dependent variables. A MANOVA univariate F analysis will allow both subscales to be simultaneously statistically analyzed, which decreased the chance of Type I error. Discrete variables assessed as appropriate for inclusion into the regression equation as a result of MANOVA were dummy coded prior to insertion into the regression equation for analysis (Tabachnik & Fidell, 2001).

Demographic and situational information was collected from both the student and instructor participants. Information collected from students included: age, gender, ethnicity, race, current grade point average, student in the accelerated or traditional track, previous experience with clinical simulations and years of experience working in health care. Information gathered from instructors included: age, gender, ethnicity, race, certification status, educational preparation and year of graduation, number of years of teaching experience and the amount of experience with simulation. Simple frequencies and percentages were calculated on this information.

Hypothesis 5a. There are no significant differences between students' mean scores on the Extent and Importance response scales between student groups who have different instructors facilitating the simulation experience.

The total summed scores for the Extent and Importance response scales were calculated for each group. A Student's *t*-test was used to assess for differences between summed mean scores on each response scale of the SPETCS. Summed scores on the SEEQ, and NCETI instruments were also analyzed for differences between student groups using the *t*-test. Differences in mean scores for individual items on all three instruments were also assessed. If significant differences ($p < .05$) are noted, group membership would be controlled for in the statistical analysis of results.

Summary

This study proposed to operationalize the concept of effective teaching in the clinical simulation environment via the newly created SPETCS instrument. The methods and procedures depicting the development of the measure and psychometric analysis to provide evidence of reliability and validity have been described. Content validity of the instrument was established through the calculation of a content validity index using the responses of an expert panel. The questionnaire was distributed to a group of undergraduate student participants in a high-fidelity clinical simulation for further psychometric testing of the scale. Internal consistency reliability was assessed through analysis of Cronbach's alpha. Test-retest reliability was examined by calculation of an ICC. Dimensionality of the instrument was evaluated through exploratory factor analysis using principle axis factoring. The eigenvalues, scree plot, and factor loadings were analyzed using both unrotated and rotated solutions with the goal of achieving simple structure. Criterion-related validity was assessed by comparing the scores on the SPETCS and a two other instruments measuring teaching effectiveness, the SEEQ and the NCTEI.

This research proposed to develop an instrument with evidence of reliability and validity to measure teaching behaviors in the context of clinical simulation.

4. RESULTS

This chapter presents the results of the development and psychometric analysis of the Student Perception of Effective Teaching in Simulation Scale (SPETCS). First, samples for the pilot test and main study are presented. Data were collected in a pilot study to assess the feasibility of the planned simulation and data collection procedures. The pilot study flowed smoothly and only minute changes to one of the instruments were made. Second, data cleaning procedures described in chapter 3 conducted prior to data analysis are discussed. Third, analysis of the data related to the specific aims and hypotheses are described.

Sample

Two convenience samples were obtained for this study, one for the pilot and the second for the main study. The pilot study was conducted in July, 2008 and the main study in October, 2008. The same site at a large Midwestern university was used for both the pilot and the main study. Twenty-nine baccalaureate nursing students were recruited to participate in the pilot and 121 students participated in the main study. Both groups were senior level nursing students enrolled in a two credit hour clinical course, Multi-System and Restorative Care, which was a required medical/surgical nursing course. One hundred percent of students enrolled in both courses agreed to complete the study instruments, however 5% (n = 6) of students in the main study did not complete the surveys due to tardiness or absence on the day of data collection. Demographic data collected included gender, age, ethnicity, race, grade point average (GPA), whether the student was in the accelerated or traditional track of the nursing program, number of previous clinical simulations, and years of experience working in healthcare. Two master teachers with experience in clinical simulations and the Multi-System and Restorative Care course served as faculty for this study. Demographic data collected from the faculty included gender, age, ethnicity, race, certification status, educational level, year of graduation, years of teaching experience, and years of teaching in simulations.

Student gender, ethnicity, race, and type of nursing program are depicted in Table 5. Both groups were very similar to one another and homogeneous. The majority of both samples were female with 89.7% (n = 26) of the pilot group and 91.7% (n = 111) of the main study group of the same gender. The ethnicity of the sample was primarily non-Hispanic or Latino, 93.1% (n = 27) in the pilot and 100% non-Hispanic or Latino (n = 121) in the larger group. One-hundred percent of the pilot sample (n = 28) and 90.1% (n = 109) of the main group were white. One student in the pilot did not indicate race. The entire pilot sample consisted of accelerated track students. The main study sample consisted of 62.8% (n = 76) traditional students and 37.2% (n = 45) accelerated track students.

Table 5
Student Gender, Ethnicity, Race, Type of Nursing Program

Characteristic	Pilot n	Pilot f(%)	Main n	Main f(%)
Gender	29		121	
Female		26 (89.7%)		111 (91.7%)
Male		3 (10.3%)		10 (8.3%)
Ethnicity	29		121	
Hispanic or Latino		2 (6.9%)		0 (0%)
Not Hispanic or Latino		27 (93.1%)		121 (100%)
Race	28		121	
American Indian or Alaska Native				1 (.8%)
Asian				4 (3.3%)
Black or African American				7 (5.8%)
White		28 (100%)		109 (90.1%)
Type of nursing program	29		121	
Traditional				76 (62.8%)
Accelerated		29 (100%)		45 (37.2%)

Participants' age, GPA, number of previous simulations, and years of experience working in health care are displayed in Table 6. The median age for the pilot group was 26 with a narrow range of 24-28 years. Median age was 23 for the main group with a much broader range than the pilot sample; 21-51 years. GPA's for the two groups were very similar, the mean and median values for the pilot group was 3.6 and those values for the main group was 3.5. The median number of previous simulations was identical for the two groups, at four with a wide range of 3-6 for the pilot and 2-8 for the main group. The median for the number of years of experience in healthcare was identical for both groups (1), but the mean for the pilot was 1.5 years with a range of 0-7 years and for the main study group the mean was 2.1 years with a range of 0-19 years.

Table 6
Student Age, Grade Point Average (GPA), Previous Simulation Experience, Years of Experience Working in Healthcare

Characteristic	n	Mean (SD)	Median	Range
Age (years)				
Pilot	29	27.3	26	24-38
Main	121	26.5 (6.4)	23	21-51
GPA				
Pilot	27	3.6 (.25)	3.6	3.0-4.0
Main	121	3.5 (.25)	3.5	2.7-4.0
Previous simulations				
Pilot	29	3.6 (.68)	4.0	3-6
Main	121	4.4 (1.3)	4.0	2-8
Healthcare experience (years)				
Pilot	28	1.5 (1.9)	1	0-7
Main	120	2.1 (3.3)	1.0	0-19

The demographic characteristics for faculty were very similar (Table 7). Both instructors taught in the course, were educationally prepared at the master's degree level, and had greater than 10 years of teaching experience. Each instructor had received certification; instructor 1 was designated as a Simulation Scholar and instructor 2 was a certified critical care nurse and certified nurse educator. One difference was evident between the number of years of simulation experience; instructor 1 had less than half the amount of experience in simulation (1 year) as instructor 2 (2.5 years).

Table 7
Instructor Demographic Characteristics

Characteristic	Instructor 1	Instructor 2
Age	52	59
Teaching experience (years)	12	18
Simulation experience (years)	1	2.5
Year of graduation	1987	1996
Gender	Female	Female
Ethnicity	Non Hispanic	Non Hispanic
Race	White	White
Certification Status	Sim Scholar*	CCRN,* CNE*

* Sim Scholar – Fairbanks Institute Simulation Scholar; CCRN – Critical Care Nursing Certification; CNE – Certified Nurse Educator

Data Screening

Data was carefully entered into the SPSS statistical software package (SPSS, Chicago, IL). After each participant's data was entered, it was double-checked for accuracy. Initially, histograms were plotted to examine the distribution of the data and identify obvious outliers. Several outliers were found in the total score of the SPETCS Extent and Importance response scales, the SEEQ and NCTEI. The data was again reexamined for accuracy and changes were made accordingly. All but 1 outlier in the SPECTS response scales were due to incorrect data entry. Correction of missing data was carefully considered, and the technique employed to rectify this issue was to take the participant's mean score on the subscale of the instrument with the missing data and input the mean value. Table 8 displays the amount of missing data in each of the instruments.

Table 8
Missing Data Percentages by Instrument

Instrument	n	*f* (%)
SPETCS A	121	3(2.4%)
SPETCS B	121	6(4.9%)
SEEQ	121	9(7.4%)
NCTEI	121	5(4.1%)

Note: SPETCS A = Extent response scale; SPETCS B = Importance response scale.

Normality of the results for each instrument was assessed through several means, with the first being visual examinations of histograms with the normal curve superimposed. Second, the skewness and kurtosis values were examined by dividing each skewness or kurtosis value by its respective standard error. The values in a normal distribution should be near zero. All of the instruments were negatively skewed, which indicated high scores. The Extent scale was skewed the least, with a *z* skew of -.29. The Importance scale had a -3.06 *z* skew value. The Extent scale was more kurtotic that the Importance scale. Lastly, a Kolomogorov-Smirnov (K-S) test was computed. The K-S values for each instrument were non-significant ($p > .05$) indicating that the results did not deviate significantly from a normal distribution with a similar mean and standard deviation (Field, 2005). Thus, data transformation was not indicated. Table 9 presents the skewness and kurtosis values and the K-S values by instrument.

Table 9
Skewness (S), Kurtosis (K), Kolomogorov-Smirnov (K-S) Test Values by Instrument

Instrument	S	K	K-S
SPETCS A	-.29	1.11	.82 ($p = .51$)
SPETCS B	-3.06	.29	1.05 ($p = .22$)
SEEQ	-1.91	-1.32	1.11 ($p = .17$)
NCTEI	-4.30	1.77	1.28 ($p = .07$)

Note: SPETCS A = Extent response scale; SPETCS B = Importance response scale.

Data screening procedures related to factor analysis such as the examination of results by instructor and the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) are presented in the portion of the chapter discussing factor analysis results. Further assessment of the normality of the results and the homoscedasticity, multicolliniarity and singularity of variables are discussed in the regression results. Next, results of the analysis based on the specific aims and hypotheses of this study are described.

Specific Aims and Hypotheses

Specific Aim 1. To develop the Student's Perception of Effective Teaching in Clinical Simulations (SPETCS) scale and determine the degree of relevance of individual items and the overall scale using Lynn's (1986) content validity index (CVI).

> *Hypothesis 1a.* The SPETCS scale items demonstrate a CVI of ≥ .78. Using a panel of nine experts, seven would need to be in agreement regarding the relevance of each item to obtain an acceptable CVI of ≥ .78 (Lynn, 1986).

Hypothesis 1a was met. Items for the SPETCS were developed based on the results of the literature review discussed in Chapter 2. A large item pool of 69 items was developed for the scale based on the 10 constructs previously identified as related to effective teaching (DeVellis, 2003). A cover letter and content validity survey (Appendices B & C) were sent to nine geographically diverse educators designated as experts in simulation by the National League for Nursing. Seven experts (78%) returned surveys for analysis, and the number of experts needed to be in agreement on each item changed to 6 out of 7 to achieve a CVI of .78 or greater. Six out of 7 in agreement equaled a CVI of .86. Reviewers were asked to rate the representativeness of each item to teaching effectiveness and indicate which construct was reflected in the item. Additional written feedback was provided by the expert panel and items were deleted or revised accordingly. Items with a CVI < .78 were examined and eliminated from the scale. From the original 69 items, the final scale used in this study contained 38 items. Each of the 10 constructs related to effective teaching were represented in the final version of the scale.

> *Hypothesis 1b.* The SPETCS total scale demonstrates a CVI of ≥ .78, which is the proportion of total items judged content valid.

Hypothesis 1b was met. A content validity index was calculated on the total scale which is the proportion of items in the scale is deemed content valid to the total number of items in the scale. Again, a CVI of .78 was the threshold as recommended by Lynn (1986). Following item revisions and deletions as recommended by the expert panel, the total scale CVI was .91.

Specific Aim 2. Evaluate the psychometric properties of the SPETCS scale.

> *Hypothesis 2a.* The SPETCS scale items demonstrate means near the center of the scale (3), the range of standard deviations indicate variability in the data, and floor and ceiling effects are < 10% (DeVellis, 2003).

Hypothesis 2a was partially met. The descriptive statistics and Cronbach's alpha for the summed scores on each response scale is presented in Table 10. The item means for the total SPETCS response scales were above 3, which is the midpoint of the 5-point scale. The overall mean for Extent response scale (A) was 4.28 and Importance response scale (B) was 4.33. The mean item standard deviations were < 1, .66 for scale A and .73 for scale B (Table 11). The standard deviation indicated a somewhat narrow range of variability in item responses. Ceiling effects for both the Extent and Importance response scales were > 10%, the mean ceiling effect for Extent was 38.4% (Table 12) and 46.6%

for Importance (Table 13). The item with the highest ceiling effect for the Extent scale (61.2%) was "the instructor was comfortable with the simulation experience," and for the Importance scale (67.8%) was "I will be better able to care for a patient with this type of problem in clinical because I participated in this simulation." The lowest ceiling percent for the Extent scale was 19.8% "I understood the objectives of the simulation," and for the Importance scale 14% "The instructor helped too much during the simulation." Hypothesis 2a was met in respect to percent floor effect for both response scales, with many items at 0%. The mean floor effect for the Extent scale was .52% and the Importance .46%. The highest percent floor for the Extent scale was 2.5% for the item "Participation in this simulation helped me to understand classroom theory," and for the Importance scale two items were at 3.3% floor "The instructor served as a role model during the simulation" and "Participation in this simulation helped me to understand classroom theory."

Table 10

SPECTS Total Scale Descriptive Statistics

Measure	n	Mean (SD)	Range	Alpha
SPETCS A	121	162.59(14.77)	119-189	.95
SPETCS B	121	164.38(18.09)	114-190	.96

Note: SPETCS = Student's Perception of Effective Teaching in Clinical Simulation Scale; A = Extent response scale; B = Importance response scale.

Table 11

Summary Item Statistics Extent and Importance Response Scales

Measure	Mean	Minimum	Maximum	Range
SPETCS A	4.28(.66)	1.88	4.59	2.71
SPETCS B	4.33(.73)	3.60	4.60	1.01

Note: SPETCS = Student's Perception of Effective Teaching in Clinical Simulation Scale; A = Extent response scale; B = Importance response scale.

Table 12

Descriptive Statistics SPETCS Extent Response Scale

Item	M (SD)	Range	% Ceiling	% Floor
Instructor help	4.12 (.586)	1-5	21.5	.80
Time allowed	4.12 (.772)	1-5	28.1	1.7
Instructor questions	4.44 (.546)	3-5	46.3	0
Autonomy	4.40 (.612)	1-5	43.8	.70
Feedback	4.50 (.534)	3-5	52.1	0
Facilitated learning	4.38 (.536)	3-5	40.5	0
Debriefing	4.53 (.518)	3-5	53.7	0
Cues	4.15 (.833)	1-5	37.2	.80
Debriefing importance	4.40 (.627)	2-5	47.1	.80
Instructor comfort	4.59 (.543)	3-5	61.2	0
Simulation interesting	4.29 (.598)	2-5	35.5	.80
Debrief questions	4.41 (.586)	2-5	44.6	.80
Realism	4.17 (.802)	1-5	34.7	.80
Understood objective	4.13 (.499)	3-5	19.8	0
Course objectives	4.45 (.512)	3-5	46.3	0
Clinical carryover	4.41 (.587)	3-5	46.3	0
Questions	4.32 (.551)	3-5	32.8	0
Develop thinking	4.37 (.534)	3-5	39.7	0
Cues	4.05 (.773)	1-5	26.4	.80
Role model	3.77 (.955)	1-5	24.8	1.7
Clinical expertise	4.30 (.762)	1-5	43.8	.80
Receptive to feedback	4.46 (.646)	2-5	52.9	0
Valuable learning	4.28 (.611)	2-5	38.0	0
Collaboration	4.47 (.578)	2-5	50.4	0
Degree of difficulty	4.30 (.669)	2-5	39.7	0
Meet expectations	4.15 (.771)	2-5	33.9	0
Cues	4.45 (.803)	1-5	32.2	1.7
Understand theory	3.95 (.874)	1-5	26.4	2.5
Collaboration	4.12 (.760)	1-5	28.9	.80
Effective strategy	4.20 (.760)	1-5	35.5	.80
Flowed smoothly	4.10 (.719)	1-5	26.4	.80
Question variety	4.31 (.545)	3-5	34.7	0
Organization	4.31 (.589)	2-5	36.4	0
Enthusiasm	4.59 (.511)	3-5	59.5	0
Cues	4.20 (.653)	2-5	32.2	0
Learning expectations	4.15 (.715)	1-5	28.1	1.7
Professional role	4.15 (.833)	1-5	36.4	.80
Questions	4.41 (.527)	3-5	43	0

Table 13

Descriptive Statistics SPETCS Importance Response Scale

Item	M (SD)	Range	% Ceiling	% Floor
Instructor help	3.59 (.881)	1-5	14	1.7
Time allowed	4.31 (.659)	3-5	42.1	0
Instructor questions	4.36 (.751)	2-5	51.2	0
Autonomy	4.38 (.733)	1-5	49.6	.8
Feedback	4.60 (.626)	2-5	66.9	0
Facilitated learning	4.46 (.684)	2-5	56.2	0
Debriefing	4.52 (.607)	3-5	57.9	0
Cues	4.26 (.750)	2-5	51.2	0
Instructor comfort	4.23 (.901)	1-5	44	1.7
Simulation interesting	4.40 (.713)	2-5	52.1	0
Debrief questions	4.30 (.691)	3-5	43	0
Realism	4.50 (.634)	2-5	51.5	0
Understood objective	4.28 (.755)	2-5	43.8	0
Course objectives	4.41 (.725)	2-5	52.9	0
Clinical carryover	4.60 (.626)	3-5	67.8	0
Questions	4.32 (.698)	3-5	45.5	0
Develop thinking	4.55 (.577)	3-5	58.7	0
Cues	4.29 (.749)	2-5	46.3	.80
Role model	3.75 (.977)	1-5	24.8	3.3
Clinical expertise	4.34 (.782)	1-5	48.8	1.7
Receptive to feedback	4.45 (.763)	1-5	57.9	.80
Valuable learning	4.41 (.691)	3-5	52.9	0
Collaboration	4.40 (.689)	3-5	51.2	0
Degree of difficulty	4.30 (.760)	2-5	46.3	0
Meet expectations	4.45 (.682)	3-5	55.4	0
Cues	4.30 (.740)	2-5	41.8	0
Understand theory	4.07 (1.04)	1-5	43	3.3
Collaboration	4.28 (.710)	1-5	40.5	.80
Effective strategy	4.32 (.721)	2-5	46.3	0
Flowed smoothly	4.23 (.728)	2-5	35.8	0
Question variety	4.09 (.876)	1-5	36.4	.80
Organization	4.43 (.669)	1-5	50.4	.80
Enthusiasm	4.45 (.741)	1-5	56.2	.80
Cues	4.33 (.712)	3-5	47.1	0
Learning expectations	4.31 (.693)	3-5	43.8	0
Professional role	4.33 (.712)	3-5	47.1	0
Questions	4.41 (.653)	3-5	49.6	0

Hypothesis 2b. Item-to-total correlations for items in the SPETCS scale are ≥ .30 among student participants in simulation (Ferketich, 1991).

Hypothesis 2b was partially met, all corrected item-to-total correlations except 1 item related to the amount of help provided by the instructor during the simulation were ≥ .30 (Tables 14 & 15). The corrected item-to-total correlation for the instructor help item in Extent scale was .07 and .24 in the Importance scale. This item was reviewed and deleted from further analysis. The reliability of the scale if individual items were deleted was reviewed to assess the impact on the reliability of the overall scale if individual items were deleted (Tables 14 & 15). The results for the Extent scale ranged between an alpha of .94 and .95 for the deletion of any given item in the scale, and the alpha for all items in Importance scale was .96.

Table 14
Item Analysis Statistics Extent Response Scale

Item	Corrected Item-to-total Correlation	Cronbach's Alpha if Deleted
Instructor help	.07	.95
Time allowed	.55	.95
Instructor questions	.62	.95
Autonomy	.60	.95
Feedback	.70	.95
Facilitated learning	.66	.95
Debriefing	.67	.95
Cues	.56	.95
Debriefing importance	.60	.95
Instructor comfort	.50	.95
Simulation interesting	.75	.94
Debrief questions	.55	.95
Realism	.50	.95
Understood objective	.46	.95
Course objectives	.55	.95
Clinical carryover	.51	.95
Questions	.62	.95
Develop thinking	.55	.95
Cues	.48	.95
Role model	.54	.95
Clinical expertise	.45	.95
Receptive to feedback	.56	.95
Valuable learning	.62	.95
Collaboration	.56	.95
Degree of difficulty	.56	.95
Meet expectations	.49	.95
Cues	.61	.95
Understand theory	.47	.95
Collaboration	.60	.95
Effective strategy	.57	.95
Flowed smoothly	.63	.95
Question variety	.54	.95
Organization	.69	.95
Enthusiasm	.57	.95
Cues	.61	.95
Learning expectations	.72	.94
Professional role	.60	.95
Questions	.60	.95

Table 15

Item Analysis Statistics Importance Response Scale

Item	Corrected Item-to-total Correlation	Cronbach's Alpha if Deleted
Instructor help	.24	.96
Time allowed	.60	.96
Instructor questions	.65	.96
Autonomy	.49	.96
Feedback	.59	.96
Facilitated learning	.72	.96
Debriefing	.66	.96
Cues	.58	.96
Debriefing importance	.63	.96
Instructor comfort	.66	.96
Simulation interesting	.59	.96
Debrief questions	.67	.96
Realism	.55	.96
Understood objective	.57	.96
Course objectives	.58	.96
Clinical carryover	.53	.96
Questions	.68	.96
Develop thinking	.65	.96
Cues	.68	.96
Role model	.60	.96
Clinical expertise	.68	.96
Receptive to feedback	.65	.96
Valuable learning	.70	.96
Collaboration	.79	.96
Degree of difficulty	.71	.96
Meet expectations	.64	.96
Cues	.61	.96
Understand theory	.56	.96
Collaboration	.66	.96
Effective strategy	.61	.96
Flowed smoothly	.69	.96
Question variety	.60	.96
Organization	.59	.96
Enthusiasm	.72	.96
Cues	.70	.96
Learning expectations	.67	.96
Professional role	.72	.96
Questions	.68	.96

Hypothesis 2c. The SPETCS scale demonstrates evidence of internal consistency reliability with Cronbach's alpha values ≥ .70 among student participants in simulation (Carmines & Zeller 1979; DeVellis, 2003; Nunnally, 1978).

Hypothesis 2c was met. The scale demonstrated internal consistency reliability as evidenced by Cronbach's alpha of .95 for the Extent scale and .96 for Importance.

Hypothesis 2d. Evidence of temporal stability of the SPETCS scale is provided by 2 week test-retest reliability with an intra-class correlation coefficient (ICC) > .60 among student participants in simulation (Shrout & Fleiss, 1979; Yen & Lo, 2002).

Hypothesis 2d was not met for the Extent scale (Table 16). The ICC was .52, which indicated moderate agreement between Time 1 and Time 2 administrations of the instrument, rather than substantial agreement between administration times. However, Pearson's r demonstrated significant correlation between administrations $r = .57$ ($p = .000$). The paired t-test result was similar to the ICC, significant differences were noted between the mean scores for Time 1 and Time 2 ($t = 3.73$, $p = .00$).

Hypothesis 2d was met for the Importance scale (Table 16). The ICC for the Importance response scale was .67, which indicated substantial agreement between participant scores. Pearson's r was equal to the ICC value at .67 ($p = .00$) and the paired t-test was not significant (p > .05) for differences in mean scores between Time 1 and Time 2. A comparison of the descriptive statistics of the Time 1 and Time 2 scores on both response scales is presented in Table 17.

Table 16
Correlations and Comparison of SPETCS Mean Scores Time 1 and Time 2

Scale	ICC (95% CI)	Pearson *r*	Paired *t* test
Extent 1			
Extent 2	.52(.37-.65)	.57**	3.7**
Importance 1			
Importance 2	.67(.55-.77)	.67**	.45

****Sig at p<.001*

Table 17
Comparison of SPETCS Scores Time 1 and Time 2

Scale	*n*	Mean(SD)	Variance
Time 1			
Extent	121	162.6(14.8)	218.2
Importance	121	164.4(18.1)	327.3
Time 2			
Extent	101	157.2(14.2)	201.3
Importance	101	163.5(18.3)	332.8

Hypothesis 2e. The dimensionality and initial evidence of construct validity of the SPETCS with factor loadings of .32 (Tabachnik & Fidell, 2001), and above for each domain is determined through exploratory factor analysis using principle axis factoring among student participants in simulation (Netemeyer, Bearden & Sharma, 2003).

Hypothesis 2e was met. Prior to factor analysis, an independent samples *t*-test was calculated to ensure that mean responses from student groups taught by different instructors were not significantly different from one another (p > .05). The means were not significantly different from one another on either the Extent or Importance subscales (Extent $t = 1.12$, $p = .27$; Importance $t = -2.56$, $p = .80$). The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) was assessed to evaluate the appropriateness of performing factor analysis (Munro, 2005). The KMO value was .88 for the Extent subscale and .91 for the Importance subscale. Both values are above the .60 minimum value recommended by Tabachnik & Fidell (2001). Bartlett's test of sphericity was significant ($p = .00$) for both subscales which indicated both correlation matrices were adequate for factor analysis.

With the support of the preliminary analyses above, initial exploratory factor analysis using principle axis factoring and varimax rotation was completed. The examination of eigenvalues > 1 produced 8 possible factors on both response scales, accounting for 70.92% of total variance for the Importance scale (Table 18) and 65.94% of the total variance for the Extent scale (Table 19). Inspection of the scree plots suggested a 3 to 4 factor solution for the Extent and a 2 to 3 factor solution for the Importance response scale.

After careful comparison of the patterns of the factor loadings with the conceptualization of effective teaching, a 2 factor solution was used for the Importance response scale: Learner Support and Real-World Application. These factors accounted for 50.60% of the variance and several items cross loaded on more than 1 factor. The values of the loadings were taken into consideration when the item loaded on both factors, and the item was placed in the factor where it loaded highest. Decisions regarding retention or elimination of individual items were based on the Importance response scale. Inter-item correlations, corrected item-to-total correlations, factor analysis and rank order of items were carefully analyzed. These results provided necessary input into participant's perceptions of which items in the scale were most important to assist achievement of simulation outcomes. As a result, a total of 4 items were deleted from the scale. Two items were related to the concept of cueing and cross loaded on both factors. The items "Cues provided support my understanding," and "Cues guided my thinking during the simulation" were deleted. These items also scored lowest of the 4 items related to cues in the rank order of means. The item "The simulation flowed smoothly" ranked poorly and had similar loadings on both factors. Lastly, the item "Instructor helped too much during the simulation" was deleted prior to factor analysis due to very poor inter-item and corrected item-to-total correlations. A forced one factor solution was completed to suggest the validity of a summed score on this instrument. The factor loadings ranged from .51-.78 which provided evidence to support the use of a summed score on the Importance response scale of the SPETCS.

Table 18

Factor Analysis of Item Pool for the SPETCS Importance Response Scale

Importance scale item	Factor 1[a]	Factor 2[b]
Instructor demonstrates clinical expertise	.79	
Instructor facilitated learning	.74	
Instructor comfort with sim	.70	
Instructor receptive to feedback	.69	
Instructor encouraged collaboration during debrief	.68	.41
Instructor encouraged collaboration during the simulation	.67	
Instructor was enthusiastic	.65	.38
Instructor as role model during sim	.63	
Cues used helped me during sim	.61	.35
Debriefing supports my understanding & reasoning	.59	.36
Instructor used variety of questions in debrief	.58	
Degree of difficulty was appropriate	.58	.44
Participation helped me understand theory	.57	
Instructor questions guided my thinking	.57	.36
Instructor lead debriefing importance	.56	.35
Questions help me better understand the situation	.55	.43
Cues provided at appropriate times	.54	.32
Sim flowed smoothly	.54	.44
Debrief questions appropriate	.53	.44
Simulation interesting	.51	.33

Table 18-*Continued*

Importance scale item	Factor 1[a]	Factor 2[b]
Cues provided supported my understanding	.51	.50
Feedback useful after sim	.47	.38
Cues provided guided my thinking	.46	.37
Sim allows me to model a professional role in a realistic manner		.81
Sims help me meet expectations when caring for real patients		.77
Sims are effective learning strategy to problem solve & make decisions		.70
Learning expectations were met		.69
Sim was a valuable learning activity	.37	.66
Sim helped develop thinking skills	.32	.64
Questions after sim helped me understand decision making	.37	.64
Time allowed to think through challenging areas of the sim		.59
Better able to care for patient with this type of problem		.58
Realism (fidelity)		.54
Sim fit with course objectives	.32	.54
Sim was well organized	.33	.53
Autonomy promotes learning		.45
Understood simulation objectives	.40	.42

[a]Factor 1 eigenvalue = 16.39, 44.29% of variance; [b]Factor 2 eigenvalue = 2.32, 6.3% of variance.

Examination of the items and their respective factors supported a 3 factor solution for the Extent response scale, which accounted for 49.26% of the total variance. Many items cross loaded onto more than one factor. Two of the 3 factors had similar patterns of factor loading and were given names identical to the Importance scale.

Factor 1 was Real-World application, Factor 2 was Debriefing/Feedback, and Factor 3 was Learner Support. A forced one factor solution was completed on this response scale to assess possible validity of a summed score on this measure. The factor loadings ranged from .48-.77 which provided evidence to support the use of a summed score on the Extent response scale. Thus, both the Importance and Extent response scales support the use of a summed score.

Table 19

Factor Analysis of Item Pool for the SPETCS Extent Response Scale

Extent Scale item	Factor 1[a]	Factor 2[b]	Factor 3[c]
Time allowed to think through challenging areas of the sim		.48	
Instructor questions guided my thinking	.62		
Autonomy promotes learning		.58	
Feedback useful after sim	.68	.33	
Instructor facilitated learning		.70	
Debriefing supports my understanding & reasoning	.59		
Cues provided guided my thinking			.75
Instructor lead debriefing importance	.42		.34
Instructor comfort with sim	.42		
Simulation interesting	.42	.61	
Debrief questions appropriate	.55		
Realism (fidelity)		.60	
Understood simulation objectives	.34	.36	
Sim fit with course objectives	.35	.40	
Better able to care for patient with this type of problem		.55	
Questions help me better understand the situation	.57	.35	
Sim helped develop thinking skills	.39	.48	
Cues used helped me during sim			.72
Instructor as role model during sim	.38		.48
Instructor demonstrates clinical expertise	.66		
Instructor receptive to feedback	.63		
Sim was a valuable learning activity		.71	

Table 19-*Continued*

Extent scale item	Factor 1[a]	Factor 2[b]	Factor 3[c]
Instructor encouraged collaboration during debrief	.50		
Degree of difficulty was appropriate		.41	
Sims help me meet expectations when caring for real patients		.73	
Cues provided at appropriate times			.75
Participation helped me understand theory	.41		
Instructor encouraged collaboration during the simulation	.39		.65
Sims are effective learning strategy problem solve & make decisions		.73	
Sim flowed smoothly	.33	.37	.45
Instructor used variety of questions in debrief	.59		
Sim was well organized	.43	.44	.34
Instructor was enthusiastic	.66		
Cues provided supported my understanding			.69
Learning expectations were met	.35	.58	
Sim allows me to model a professional role in a realistic manner		.63	
Questions after sim helped me understand decision making	.34	.46	

[a]Factor 1 eigenvalue = 13.90, 36.58% of variance; [b]Factor 2 eigenvalue = 2.77, 7.28% of variance; [c]Factor 3 eigenvalue = 2.05, 5.40% of variance.

After careful consideration of the above results, several items were deleted from the scale. The item "The instructor helped too much during the simulation" was deleted prior to factor analysis due to poor results on the item analysis. In addition, this item had the lowest mean score on the Importance response scale. Four additional items were deleted based on the factor loadings. Two items specifically related to cues, “Cues (hints) provided during the simulation guided my thinking” and “Cues provided during the simulation supported my understanding” were deleted based on factor loadings of .46 and .37 for the former and .51 and .50 for the latter. The two additional items related to cueing were retained in the scale based their factor scores and mean score rank, “Cues were provided at appropriate times during the simulation” and “Cues were used in the simulation to help me progress through the experience.” The first item loaded on the factor 1 only (.53) and the second item loaded on both factors, but loaded much higher on factor 1 (.58) than on factor 2 (.34).

Two additional items were deleted, “I understood the objectives of the simulation” and “The clinical simulation flowed smoothly.” Both loaded on Factors 1 and 2, with similar factor loading values. The item related to objectives scored .40 on Factor 1 and .42 on Factor 2. The item related to the flow of the simulation loaded at .54 on Factor 1 and .44 on Factor 2. The mean score rank for these items was in the bottom seven out of 38 in the original item pool.

Following deletion of these five items, factor analysis was completed on the items that remained (Table 20). This analysis with a two factor solution accounted for 51.57% of the total variance, which was a slight improvement from the 50.60% in the original analysis. No additional items were deleted from the scale. The result was 20 items that loaded primarily on Factor 1 and 13 items loaded on Factor 2. Items that cross-loaded on both factors were reviewed to examine the difference between the loadings, rank mean scores on the scale, item analysis results and the relationship of the loading pattern to the conceptual framework for this study.

Table 20
Factor Analysis of the 33 item SPETCS Importance Response Scale

Item	Factor 1	Factor 2
1. Instructor clinical expertise	.80	
2. Instructor facilitated learning	.74	
3. Instructor comfortable during sim	.70	
4. Instructor receptive to feedback	.69	
5. Encouraged collaboration in debrief	.67	.41
6. Enc. collaboration during simulation	.67	
7. Instructor enthusiastic	.66	.38
8. Instructor as role model	.62	
9. Debrief supports reasoning	.59	.37
10. Difficulty appropriate	.59	.44
11. Question variety in debrief	.58	
12. Cues helped progression through sim	.58	.34
13. Questions after sim guide thinking	.57	.37
14. Better understand theory	.57	
15. Questions help understand situation	.58	.43
16. Importance of debriefing	.55	.35
17. Cues provided at appropriate times	.53	
18. Questions appropriate during debrief	.52	.44
19. Simulation interesting	.50	.33
20. Instructor provided useful feedback	.47	.39
21. Sim allows role modeling		.81
22. Sim helps meet clinical expectations		.77
23. Effective learning strategy		.70
24. Learning expect met		.69
25. Participation valuable learning	.37	.66
26. Sim develops critical thinking	.33	.65
27. Questions help decision-making	.37	.65
28. Allowed time to think in sim		.60
29. Better able to care for patient in clinical		.58
30. Simulation fidelity		.54
31. Simulation was organized	.32	.54
32. Sim objectives fit with course		.53
33. Instructor provides autonomy		.46

The new 33 item scale was further analyzed to examine for internal consistency reliability of the entire scale and the subscales of Learner Support and Real-world Application in the same fashion as the original scale. Alphas were acceptable and ranged from .92-.96 (Table 21).

Table 21
Reliability of 33 item SPETCS Extent and Importance Response Scales and Subscales using Cronbach's Alpha

Measure	*n*	items	Alpha
Extent	121	33	.94
Importance	121	33	.96
Learner Support	121	20	.95
Real-world Application	121	13	.92

Hypothesis 2f. The SPETCS demonstrates evidence of criterion-related validity as evidenced by significant ($p < .05$) correlation with the Student Evaluation of Educational Quality (SEEQ, Marsh, 1987) and the Nursing Clinical Teacher Effectiveness Instrument (NCTEI; Mogan & Knox, 1987).

Hypothesis 2f was met. The SPETCS demonstrated evidence of criterion-related validity. Data from the Extent of Agreement and Importance response scales and the SEEQ were pooled as the mean scores on those instruments between student groups with instructor 1 and instructor 2 were not significantly different (Table 22). However, mean scores between student groups were significantly different ($p < .05$) between instructors on the NCTEI and the correlations comparing the NCTEI and the Extent and Importance scales were compared separately, by instructor. All correlations were significant ($p < .01$), which supported the instrument as criterion valid. Results of the correlations are displayed in Table 23.

Table 22

Differences in Mean Scores by Instructor

Instrument	*n*	Mean(SD)	*t* value
Extent			
Instructor 1	62	161.8(13.7)	
Instructor 2	59	158.8(15.7)	1.12
Importance			
Instructor 1	62	164.0(19.0)	
Instructor 2	59	164.8(17.2)	-.26
SEEQ			
Instructor 1	62	129(11.2)	
Instructor 2	59	129(17.3)	-.01
NCTEI			
Instructor 1	62	303.8(22.5)	
Instructor 2	59	292.9(29.9)	2.28*

* $p < .05$

Table 23

Correlation between SPECTS Extent and Importance Response Scales and Criterion Instruments using Pearson's r

Instrument	*n*	Pearson's *r*
Extent		
SEEQ	121	.49**
Importance		
SEEQ	121	.40**
Extent		
NCTEI		
Instructor 1	62	.52**
Extent		
NCTEI		
Instructor 2	59	.57**
Importance		
NCTEI		
Instructor 1	62	.52**
Importance		
NCTEI		
Instructor 2	59	.39**

** $p < .01$

Specific Aim 3. Determine which teaching strategies/behaviors are most frequently used in clinical simulations based on perceptions of student participants.

The items on the Extent of Agreement response scale were rank ordered based on the mean scores. Instructor strategies/behaviors used to the greatest extent as perceived by participants are listed in Table 24. The results suggested that the instructors: 1) were comfortable with the simulation process, 2) were enthusiastic during simulation, 3) provided useful feedback, 4) encouraged collaboration and 5) supported student clinical reasoning during debriefing.

Table 24
Rank Order of Items by Mean Score (SD) on Extent Response Scale

Item	Mean(SD)
1. Instructor comfort with simulation	4.59(.54)
2. Enthusiasm	4.59(.51)
3. Debriefing supports reasoning	4.53(.52)
4. Feedback useful	4.50(.53)
5. Collaboration encouraged in debrief	4.47(.58)
6. Instructor receptive to feedback	4.46(.65)
7. Simulation fit course objectives	4.45(.53)
8. Questions guide thinking	4.44(.55)
9. Debriefing importance	4.41(.63)
10. Debrief questions appropriate	4.41(.59)
11. Improved clinical ability	4.41(.59)
12. Questions help decision making	4.41(.53)
13. Autonomy promotes learning	4.40(.61)
14. Instructor facilitates learning	4.38(.54)
15. Simulation developed critical thinking	4.37(.53)
16. Questions help situational understanding	4.32(.55)
17. Instructor clinical expertise	4.31(.76)
18. Degree of difficulty appropriate	4.31.(67)
19. Simulation was organized	4.31(.59)
20. Question variety in debrief	4.31(.55)
21. Simulation interesting	4.29(.60)
22. Simulation valuable learning tool	4.28(.66)
23. Simulations help problem solving	4.20(.76)
24. Cues support understanding	4.20(.65)
25. Simulation fidelity	4.17(.80)
26. Simulation allows professional role modeling	4.15(.83)

Table 24-*Continued*

Item	Mean(SD)
27. Cues guide thinking	4.15(.83)
28. Cues at appropriate times	4.15(.80)
29. Meets expectations with real patients	4.15(.77)
30. Learning expectations met	4.15(.71)
31. Understand simulation objectives	4.13(.50)
32. Collaboration encouraged in simulation	4.12(.76)
33. Instructor as role model	4.12(.59)
34. Time allowed to think during simulation	4.11(.77)
35. Simulation flowed smoothly	4.09(.72)
36. Cues were helpful	4.05(.77)
37. Improved understanding of theory	3.95(.87)
38. Instructor as role model	3.77(.96)

Specific Aim 4. Determine which teaching strategies/behaviors best facilitate achievement of specified simulation outcomes based on ratings of student participants in simulation.

Items on the Importance response scale were rank ordered based on mean scores to determine student perceptions of the strategies/behaviors which most facilitated the achievement of the simulation objectives (Table 25). Items with the 5 highest means included: 1) useful feedback from the instructor, 2) improved ability of the participant to care for a patient with similar health problems in clinical as a result of participation in simulation, 3) the development of critical thinking abilities during the simulation experience, 4) perception that the debriefing component of the simulation experience supported clinical reasoning and 5) the importance of the fidelity of the simulation. Two items were in the top five for both response scales: useful feedback and debriefing supports reasoning. Three items were in the bottom five items for each response scale: 1) instructor as a role model, 2) the simulation flowed smoothly and 3) the simulation improved understanding of theory.

Table 25

Rank Order of Items by Mean Score (SD) on Importance Response Scale

Item	Mean (SD)
1. Useful feedback	4.60(.63)
2. Improved clinical ability	4.60(.63)
3. Simulation developed critical thinking	4.55(.58)
4. Debriefing supports reasoning	4.52(.61)
5. Simulation fidelity	4.50(.63)
6. Instructor facilitates learning	4.46(.68)
7. Instructor receptive to feedback	4.45(.76)
8. Instructor enthusiasm	4.45(.74)
9. Meets expectations with real patients	4.45(.68)
10. Simulation was organized	4.43(.67)
11. Simulation fit course objectives	4.41(.73)
12. Simulation valuable learning tool	4.41(.69)
13. Questions help decision making	4.41(.65)
14. Simulation interesting	4.40(.71)
15. Collaboration in debriefing	4.40(.69)
16. Autonomy promotes learning	4.38(.73)
17. Questions guide thinking	4.36(.75)
18. Debriefing importance	4.36(.36)
19. Instructor clinical expertise	4.35(.78)
20. Cues support understanding	4.33(.71)
21. Simulation allows professional role modeling	4.33(.71)
22. Simulations help problem solving	4.32(.72)
23. Questions help situational understanding	4.32(.70)
24. Cues at appropriate times	4.31(.74)
25. Learning expectations met	4.31(.69)
26. Time allowed to think in simulation	4.31(.66)
27. Degree of difficulty appropriate	4.30(.76)
28. Cues were helpful	4.30(.75)
29. Debrief questions appropriate	4.30(.69)
30. Understood simulation objectives	4.28(.76)
31. Collaboration during simulation	4.28(.71)
32. Cues guided thinking	4.26(.75)
33. Simulation flowed smoothly	4.23(.73)
34. Instructor comfort with simulation	4.23(.90)
35. Question variety in debriefing	4.09(.88)
36. Improved understanding of theory	4.07(1.04)
37. Instructor as role model	3.75(.98)
38. Instructor assistance	3.56(.88)

Specific Aim 5. Determine the relationship between student participant's demographic and situational variables and their responses on the Extent and Importance response scales of the SPETCS. Determine if scores between the two student groups who had different instructors are similar.

Determination of the relationship between demographic and situational variables to participant scores on the two response scales was assessed with regression analysis. Prior to entering independent variables into the regression equation, variables were screened to determine if there was a significant correlation between total scores on the scales and each variable. The continuous variables of student's age, GPA, previous simulation experience and experience working in healthcare settings were screened using Pearson's correlation coefficient (Table 26). Two variables were significantly correlated with a response scale. Age was significantly correlated with the Importance response scale (r = .19, p < .05), which indicated that as age increased the total score on the Importance response scale increased. The amount of previous work experience in health care was significantly correlated with the Extent of Agreement response scale (r =.22, p *<.05).* Those participants with more healthcare work experience scored higher on the Extent of Agreement response scale.

Table 26
Correlations of Continuous Demographic and Situational Variables to Screen for Inclusion in Regression Analysis using Pearson's r

Variable	Response Scale	
	Extent	Importance
Age	.11	.19*
GPA**	-.11	.06
Sim Experience	-.04	-.03
Work Experience	.22*	.08

*p<.05; **Grade Point Average

Age was entered into a simple linear regression equation to further examine the relationship between age and the total score on the Importance response scale. A significant regression equation was found [F (1,119) = 4.40, p = .038], with an R^2 of .036. Thus, 3.6% of the variance in scores was accounted for by age. Work experience in healthcare was entered into a simple linear regression equation to evaluate the relationship between work experience and the Extent of Agreement response scale. This regression was significant [F (1,118) = 5.79, p =.018], with an R^2 of .047. The result indicated that 4.7% of the variance in scores on the Extent of Agreement scale was accounted for by participant's health care work experience.

Next, the discrete variables of gender, accelerated or traditional program track and race were screened for inclusion into regression analysis using MANOVA univariate F (Table 27). None of the F values were significant (p > .05). As a result, none of these discrete variables contributed significantly to the variance in the total scores on each response scale and were not analyzed with regression.

Table 27
Screening of Discrete Demographic and Situational Variables for Inclusion in Regression Analysis using MANOVA univariate F

Variable	Response Scale	
	Extent	Importance
Gender[a]	.05	.37
Program track[b]	1.32	1.49
Race[c]	.15	.97

[a]*df(1,119);* [b]*df(2,118);* [c]*df(1,119)*

Hypothesis 5a. There is no significant difference between students mean scores on the Extent and Importance response scales between student groups who have different instructors facilitating the simulation experience.

Hypothesis 5a was met. Mean scores on the Extent and Importance subscales were not significantly different between student groups based on instructor (Extent $t = 1.12$, $p = .27$; Importance $t = -2.56$, $p = .80$). Thus, data from both student groups was pooled for analysis.

Summary

This chapter described the data cleaning and analysis procedures used to evaluate the psychometric properties of the SPETCS. The original scale contained 59 items and was evaluated by an expert panel. After revisions to the scale were completed based on content expert feedback, the new 38 item scale had evidence of content validity as demonstrated by a CVI of .91. A convenience sample ($n = 121$) of senior baccalaureate nursing students completed the study instruments. Descriptive statistics and item analyses were analyzed on all items to evaluate for individual item retention or deletion. The instrument was found to have evidence of internal consistency reliability on both the Extent ($\alpha = .95$) and Importance ($\alpha = .96$) response scales. Temporal stability reliability was assessed by an intra class correlation coefficient. The ICC for the Extent scale was lower than expected at .52, and for the Importance scale the ICC = .67 which met expectations. The summed scores on the SPETCS were compared with two well-known instruments to assess criterion-related validity. Pearson's r compared results from both both response scales with both criterion instruments. Correlations were moderate, with a range from .39-.57.

Exploratory factor analysis using principal axis factoring was conducted to assess the dimensionality of the scale and provide additional insight into the retention and deletion of items. The EFA results from the Importance response scale were used to analyze items. Two factors were found in the Importance scale which accounted for 51.57% variance: Learner Support and Real-World Application. The Extent scale had three factors accounting for 49.26% variance: Learner Support, Feedback/Questioning and Real-World Application. The mean rank scores for individual items on the Importance scale were also used to make decisions with regard to item retention and deletion. Simple linear regression examined the relationship between demographic and situational characteristics of participants. A relationship was found between age and the importance response scale and between the amount of healthcare work experience and the Extent response scale. The result of the instrument development and psychometric testing of the SPETCS was a 33 item instrument with items from each of the 10 concepts identified in Chapter 2. Discussion of the results related to the literature, research and practice implications are presented in Chapter 5.

5. DISCUSSION

This chapter presents a discussion of the psychometric analysis of the Student Perception of Teaching Effectiveness in Clinical Simulation scale based on the specific aims of this study. The limitations of the study put forth in Chapter 1 are addressed. Lastly, the theoretical, research and practice implications for nurse educators are described.

Specific Aims

Specific Aim 1. To develop the Student Perception of Effective Teaching in Clinical Simulation (SPETCS) scale and determine the degree of relevance of individual items and the overall scale using Lynn's (1986) content validity index (CVI).

Specific Aim 1 was met as the results of the content validity assessment were within the hypothesized parameters (Lynn, 1986). Recommendations were followed to include a large item pool in early drafts of a new instrument (DeVellis, 2003). Several items in the original pool were very similar; the expert feedback and a CVI value was calculated on each item provided guidance to select the best items for the psychometric evaluation of the instrument. The original draft of the instrument contained 69 items and as a result of the content validity assessment, a scale with 38 items emerged for continued analysis. The overall CVI for the 38 item revised scale (.91) met expectations. Items representative of each of the constructs identified from the literature review were included.

The selection of geographically diverse and nationally recognized content experts (n = 7) for input related to content validity was a notable strength of this study. No other studies found in the education literature described how experts were selected or used content experts of this quality. In addition, none of the research quantified content validity using Lynn's (1986) criteria. The use of these rigorous methods to ensure this instrument has credible evidence of content validity is a unique feature of this study.

Specific Aim 2. Evaluate the psychometric properties of the SPETCS scale.

Mean item scores for both the Extent and Importance response scales were above the midpoint (4.28; 4.33). The ceiling effects were > 10%; the mean ceiling effect for the Extent response scale was 38.44% and the Importance scale was 46.58%. The higher than expected average mean scores and ceiling effects of items on the Extent response scale were likely due to the fact that the instructors (*n* = 2) who facilitated the simulation experience were master teachers with significant general teaching experience and experience in clinical simulation. One of the instructors was a Fairbanks Institute Simulation Scholar, and the other held multiple certifications in clinical specialty areas and nursing education. In addition, a state-of-the-art, high-fidelity simulation laboratory was the site of the study. Thus, it would be expected that these experienced instructors would use effective teaching behaviors in such a quality laboratory environment, and as a result, receive higher than average scores on the Extent response scale.

The average mean and ceiling effect scores on the Importance response scale were higher than on the Extent scale. The rationale for the high values was the use of the NESF model to design the simulation used for the study. This simulation was carefully designed to follow specific guidelines developed based on the model. These results further validate the findings of earlier studies of the development and implementation of simulations which used the NSEF model as a framework (Jeffries, 2006). Student perceptions of the overall simulation experience were very positive; teaching strategies and behaviors used in this simulation supported attainment of the learning objectives.

Item analysis was carefully considered using Ferketich's criteria (1991). All items in both subscales met specified criteria for inclusion except one item that clearly did not perform well. This item was eliminated early in the analysis. The alpha if item deleted results indicated that if an individual item were deleted from the either response scale, the internal consistency of the overall scale would continue to be satisfactory (> .70). The alpha if deleted ranged from .94-.96 for individual items.

The assessment of internal consistency reliability for the Extent and Importance response scale met the hypothesized expectations. The alpha value for the Extent scale was equal to .95 and alpha was .96 for the importance subscale. These values are well above the .70 threshold that is commonly cited in the psychometric literature (Carmines & Zeller 1979; DeVellis, 2003; Nunnally, 1978). Coefficient alpha quantifies the "degree of interrelatedness among a set of items" (Netemeyer, Bearden & Sharma, 2003, p. 49). This result provides evidence that the SPETCS items are strongly related to the dependent variable of effective teaching due to the fact that they are strongly related to one another. The length of the scale impacts the alpha value, but at this stage of instrument development, it is necessary to retain those 33 items to sufficiently represent the 9 constructs that relate to effective teaching.

The intra-class correlation used to assess temporal stability for the Importance response scale met hypothesized expectations (ICC = .67). This result indicated substantial agreement between participant responses at the original and subsequent administrations of the instrument. Thus, student perceptions of the importance of the teaching behaviors and strategies used in the simulation toward meeting the simulation objectives did not change in the approximate 2 week interval between administrations.

The intra-class correlation for the Extent response scale did not meet hypothesized expectations (ICC = .52; 95% CI .37-.65). The Pearson's r ($r = .57$; $p < .001$) was significant, indicative of a moderate correlation between administrations (Kerlinger & Lee, 2000). However, a *t*-test indicated significant differences between administrations ($t = 3.7$; $p < .05$). Mean scores on the first administration of the measure were 5 points higher than on the second administration. This finding was unexpected; in the literature, effective teaching was considered to be a stable trait at over a short time period, such as the two week retest time (D'Apollonia & Abrami, 1997; Marsh, 1987).

Several possible issues exist that could have contributed to the lower than expected ICC on the Extent response scale. First, the effect of history may have impacted results. Participants had contact with both of the instructors who conducted the simulations between time one and time two which may have impacted responses on the time two administration. Second, students may have had examinations in this or another course between time one and two that may have impacted results. Third, there was variation in the amount of time elapsed between administrations of the instrument. Clinical instructors were asked to administer the SPETCS approximately 2 weeks after the simulation during clinical post-conference. There was variation between groups; with a range of 1-3 weeks. Post-conference locations where the questionnaires were distributed were different which may have affected results. As a result, further assessment of the temporal stability of the Extent response scale was identified as an area for future research, with a focus on standardization of the retest time interval and setting.

Assessment of the dimensionality of the SPETCS early in the instrument development process was the primary purpose of exploratory factor analysis (Netemeyer, Bearden & Sharma, 2003). EFA also provides key information to support decisions related to item retention and deletion The Importance response scale was chosen over the Extent of Agreement response scale for factor analysis of the SPECTS. The rationale for the selection of the Importance scale was that participant perceptions (*n* = 121) of the importance of individual teaching strategies and behaviors toward meeting simulation objectives provided more valuable insight than the evaluation of two master teachers on the Extent of Agreement scale at this stage of instrument development.

Two factors, Learner Support and Real-World Application emerged from the factor analysis; these factors accounted for 51.57 % of the variance. Learner Support factor loadings ranged between .46-.79 and loadings on the Real-World Application factor ranged between.42-.81. Several items cross-loaded on both factor and decisions about individual items were made based on loading patterns on each factor, the literature and the individual item rank order on the scale (Specific Aim 4). As a result, four items with small differences between factor loadings on each factor on were deleted from the original scale (Table 20). The final product was a 33 item scale which included items based on each of the constructs identified in the literature.

The SPETCS demonstrated evidence of criterion-related validity with significant ($p < .01$) Pearson product-moment correlations between each response scale and the two criterion instruments, SEEQ and NCTEI (Table 16). Prior to pooling data from participants with different instructors, the scores on all three instruments were assessed to ensure there were no differences between mean scores by instructor using a Student's *t*-test (Table 17). An unexpected, significant difference between mean scores on the NCTEI by instructor was found ($t = 2.28$; $p < .01$). No differences were found between mean scores by instructor on both response scales of the SPETCS and the SEEQ. The demographic and situational variables of each instructor were very similar; both were considered to be master teachers in simulation. The only difference noted was that instructor 1 was a Fairbanks Simulation Scholar and instructor 2 was not. The Fairbanks Simulation training consisted of a week-long immersion in simulation theory, development and implementation. This may have had an impact on those results. Correlations between the NCTEI and SPETCS response scales by instructor met expectations; moderate correlations were found for both Instructor 1 ($r = .52$; .52) and Instructor 2 ($r = .57$; .39).

Specific Aim 3. Determine which teaching strategies/behaviors are most frequently used in clinical simulations based on perceptions of student participants.

Participant mean scores on Extent of Agreement response scale items were rank ordered to assess which teaching strategies/behaviors were most frequently used in the simulation (Table 24). Items with the highest mean scores included: Instructor comfort with simulation, instructor enthusiasm, debriefing, feedback, and collaboration. Items with the lowest mean scores included: time to think during the simulation, simulation flowed smoothly, helpful cues, improved understanding of theory, and instructor as a role model.

Specific Aim 4. Determine which teaching strategies/behaviors best facilitate achievement of specified simulation outcomes based on ratings of student participants in simulation.

The determination of which teaching behaviors best facilitated attainment of simulation outcomes was based on the mean score rankings on the Importance response scale. Behaviors perceived to be most important by participants included: feedback, behaviors which improved clinical ability and critical thinking, the support of clinical reasoning through the debriefing process, and the fidelity of the simulation. Behaviors which scored lowest in the rankings included: instructor comfort with the simulation, the variety of questions used in the debriefing, improved understanding of theory from the simulation, the instructor as a role model, and instructor assistance during the simulation. In addition, these rankings were taken into consideration during the item and factor analysis.

These findings are similar to results from the National League for Nursing/Laerdal studies (Jeffries & Rizzolo, 2006), which found feedback, debriefing, and fidelity as critical design elements of a successful simulation. The results also support the importance suggested simulation outcomes from the NESF: critical thinking and clinical reasoning. Student participants clearly indicated the importance of the real-world application of what was learned during the simulation to the clinical setting.

Specific Aim 5. Determine the relationship between student participant's demographic and situational variables and their responses on the Extent of Agreement and Importance response scales of the SPETCS. Determine if scores between the two student groups who had different instructors are similar.

Regression analysis was used to determine if a relationship existed between the demographic and situational variables of participants and their scores on the SPETCS. Prior to inclusion in the regression equation, data was screened to find variables that were significantly related to SPETCS scores. Continuous variables were screened using Pearson's product moment correlation coefficient and discrete variables were screened using MANOVA univariate *F*. Only 2 variables were significant, age was positively correlated with the Importance response scale ($r = .19$; $p < .05$) and the number of years of healthcare experience was positively correlated with the Extent of Agreement scale ($r = .22$; $p < .05$). Each variable was entered into a linear regression equation to quantify the relationship; 3.6% of the variance in the score on the Importance scale could be attributed to age. The older participants had higher mean scores. The amount of healthcare experience accounted for 4.7% of the variance in the Extent scale. Participants with more healthcare experience had higher mean scores on the Extent scale. The relationship between these variables and results on the SPETCS is an area for future research.

Theoretical Implications

Clinical simulation is a unique teaching/learning strategy which requires similar but not identical teaching strategies to those used in didactic and clinical contexts to meet learning outcomes. Well designed simulations are a learner-centered, active strategy requiring the learner to apply, evaluate and synthesize knowledge in the affective, cognitive and psychomotor domains in a safe, controlled environment.

The results of this study provide insight into the role of the teacher and effective teaching behaviors in clinical simulation. The role of the teacher and teaching effectiveness within this specialized educational context had not been previously defined or studied empirically and was identified as a significant gap in the literature. The Nursing Education Simulation Framework (NESF) was the conceptual model used to design the simulation and guide development of the Student Perception of Effective Teaching in Clinical Simulation scale (SPETCS).

The SPETCS was developed as a means to examine the role of the teacher and evaluate teaching behaviors empirically within simulation contexts. The instrument was developed based on the tenets of socio-cultural, constructivist and learner-centered educational theories which underpin the NESF. In addition, related literature from classroom and clinical education arenas assisted with the identification of the 10 constructs used to guide the development of the SPETCS.

The exploratory factor analysis of the importance response scale of the SPETCS revealed two factors, Learner-Support and Real-World Application. These factors are more global in nature as the 10 original constructs were collapsed into two (Table 28). Learner-Support encompassed items which appeared to fit with learner-centered and constructivist learning theories. Those items related to what occurred during the simulation experience such as enthusiasm, feedback, cues, questioning and the facilitation of learning. Real-World Application items related heavily to the expectations of the learner as a result of participation in the simulation. The simulation fidelity, the ability to better care for a client in the clinical setting as a result of participation, the development of critical thinking abilities and that learning expectations were met are examples of items on this factor that support the tenets of socio-cultural and constructivist learning theory.

Table 28
Constructs Underlying Factors in Importance Response Scale

Factor	Construct
Learner-Support	Feedback and Debriefing
	Teaching Ability
	Modeling*
	Interpersonal Relationships
	Enthusiasm
	Cuing
	Questioning
Real-World Application	Expectations
	Organization
	Modeling*

**Related items loaded on both factors*

The two factors suggested in the results of this study fit with the multiple dimensions identified in the teaching effectiveness literature from the clinical and classroom settings (Tables 2 & 4). Most instruments had between two and five dimensions. This body of literature proposed that these multidimensional instruments would provide more guidance to faculty than a general one-dimensional measure of effective teaching. These instruments highlight specific areas of strength as well as areas where improvement may be needed. The SPETCS, with two dimensions and response scales, has the potential to provide pertinent feedback to faculty using clinical simulations in the teaching/learning process.

Interestingly, the construct of feedback was common to several studies reviewed in Chapter 2 (Copeland & Hewson, 2000; Kirshling et al., 1995). The NLN-Laerdal simulation studies, from which the NSEF was developed, highlighted the importance of the debriefing component of the simulation to promote student learning (Jeffries, 2005; Jeffries & Rizzolo, 2006). Both the non-empirical and research based simulation literature consistently highlight debriefing and feedback from the instructor as one of the prominent features of a successful clinical simulation.

Research Implications

This study was a first step in the creation of an instrument to assess teaching effectiveness in simulations. Future research needs to be considered that further examines the psychometric properties of the SPETCS. The sample population in this study was very homogeneous, and the administration of this measure to more diverse student groups in varied types of nursing educational programs is necessary to support generalization to other student populations. The temporal stability of the Extent of Agreement response scale requires further study due to the lower than expected intra-class correlation results between administrations of the instrument. The instructors involved in the study were highly-qualified master teachers in simulation contexts. Future studies should include faculty with varied degrees of experience; one of the primary aims of the study was to develop an instrument which would promote faculty development through the identification of areas of strength and areas for improvement.

The creation of context-specific evaluation tools for educators has been recommended in the literature (D'Appolonia & Abrami, 1997). In addition, this research addressed several of the problematic methodological issues in the research related to clinical simulations (Bradley, 2006; Seropian, 2003). The paucity of research in the area of effective teaching in simulations coupled with empirical studies that were primarily atheoretical, with poor methodological rigor, small sample sizes, and measures without evidence of reliability and validity were problematic. This study addressed those issues and can serve as a model to guide future empirical studies in education.

The SPETCS was created using principles outlined in the instrument development literature (DeVellis, 2003; Netemeyer, Bearden & Sharma, 2003). The instrument was developed using a theoretical framework grounded in established learning theory. Conceptual and operational definitions were carefully considered and clearly written. Items were developed based on constructs identified in the related literature and followed item writing guidelines (Dillman, 2000).

Established criteria were used to begin assessment of reliability and validity (Carmines & Zeller, 1979; DeVellis, 2003). The first step was to assess content validity. Nationally known experts reviewed the instrument and their input was analyzed using Lynn's (1986) criteria. A content validity index was computed and decisions were made about the retention and wording of items. The initial draft of the instrument contained a large pool of items, and the CVI guided decisions to reduce item numbers to a manageable amount. Quantification of content validity was a necessary first step to provide evidence of the validity of the instrument. Internal consistency reliability was assessed with the well-known Cronbach's alpha. The temporal stability of the instrument was assessed using an intra-class correlation coefficient, which recent psychometric literature suggested was more appropriate than a Pearson's *r* to assess the correlation between administrations of the instrument (Yen & Lo, 2002).

Item analysis was completed using Ferketich's criteria (1991) criteria, and dimensionality assessed through exploratory factor analysis with principle axis factoring. Principle axis factoring has been purported to be an improvement over more principle components analysis when using factor analysis to assist with instrument development (Netemeyer, Bearden & Sharma, 2003). Sample sizes necessary for factor analysis must be sufficient, and the psychometric theory literature must be consulted. In addition to basic guidelines for sample size, statistical procedures to assess sampling adequacy are available.

The developer of the instrument must understand and use their knowledge of the theoretical basis underpinning the instrument throughout the process; statistics alone are insufficient. Decisions made in regard to the instrument must make conceptual sense. Careful consideration of the theory and conceptual definitions developed early in instrument development process cannot be overstressed. In sum, one of the major contributions of this study relate to the application of measurement theory to the instrument development process for use in nursing education.

Practice Implications for Nurse Educators

Nurse educators are charged to develop and use evidence-based educational practices to prepare graduates able to function in the complex healthcare environment (NLN, 2005). The SPETCS was created and psychometrically analyzed to begin to define and evaluate teaching strategies within simulation contexts. This easy to administer, 33 item instrument has the potential to provide instructors with theory-based guidance in the development stage and learner feedback in the evaluative stage of the implementation of clinical simulations.

These results propose a set of teaching behaviors/strategies that learners perceive to be most important to the attainment of simulation outcomes. It is suggested that these items be taken into careful consideration when a simulation is planned. Most important to the learners in this study was the provision of useful feedback from the instructor. In most well-designed simulations, feedback is generally provided during the debriefing immediately following the simulation. However, feedback may be given at any time depending upon the simulation's purpose. Another noteworthy finding was the importance of the debriefing to assist learner development of clinical reasoning abilities. Learners should be encouraged to share and discuss the thinking which guided their decisions made during simulation. Educators need to thoughtfully plan the debriefing as these results add additional support to previous studies of the high quality learning which occurs during this time.

Another area identified as important to learners was the capability to translate learning from the simulation laboratory into improved ability to care for actual patients in the clinical setting. These results imply that the creation of simulations that are realistic and allow learners to model psychomotor, cognitive and affective behaviors needed in the clinical area are necessary components of well-designed simulations. The fidelity of the simulation may be enhanced through the use of high fidelity human patient simulators, in a realistic setting. Other suggestions include clinical uniforms worn by both the instructor and learners, and the use of realistic props. These results highlight the need for nurse educators to carefully develop simulations with particular attention paid to areas identified as important to learners as they are integrated into the curriculum.

Limitations

The limitations of the study outlined in the Introduction will be discussed. Each limitation is presented individually and related to the research findings.

1. A non-probability, homogeneous, convenience sample will be used in this study.

A non-probability convenience sample of 121 senior level nursing students from one baccalaureate program served as the sample for the study (Table 5). A majority of the participants were Caucasian (90.1%) and female (91.7%). The homogeneity of the sample was expected, and this limitation was considered to be acceptable at this stage in the development of a new instrument. Further evaluation of the SPECTS in other types of nursing programs with more heterogeneous student populations is a necessary next step to assess the generalizability of these findings.

2. Characteristics of effective teaching will be measured by a newly developed instrument with no prior psychometric analysis.

The Nursing Education Simulation Framework was the comprehensive conceptual model that directed the instrument development and the clinical simulation used this study. The SPETCS was developed with careful attention to the socio-cultural, constructivist and learner-centered theories of education that underpin the NESF. The current literature related to effective teaching from both classroom and clinical settings was also integral to the development of this instrument. Although the SPECTS is a new instrument, the comprehensive theory and literature which guided instrument development is clearly a strength of this study.

The psychometric analysis of the SPETCS was rigorous. Guidelines for scale development and psychometric analysis were grounded in the literature. Great care was taken to follow established guidelines as closely as possible and clearly delineate procedures to allow for replication of the study, which has been a weakness in educational research (Reed et al., 2005).

Particular attention needs to be paid to the assessment of the temporal stability of the instrument related to scores on the Extent of Agreement response scale of the SPETCS. Scores on between time 1 and time 2 administrations did not correlate as well as expected. Standardization of the timing and setting of the retest needs to be considered in future research.

3. A cross-sectional, descriptive design will be used.

The cross-sectional, descriptive design was appropriate for the purpose of this research, which was to develop a new instrument. In future studies, a longitudinal design may provide insight into changes related to what learners perceive as important strategies and behaviors. In addition, over time it could be expected that an instructor with little to no experience facilitating clinical simulations would receive higher summed scores on the Extent of Agreement response scale.

Recommendations for Future Research

A significant gap in the literature exists related to the identification of best practices related to the role of the teacher and effective teaching strategies/behaviors specifically in simulation contexts. This study was designed to begin to address this gap and create a reliable and valid tool to assist nurse educators to improve the quality of the simulation experience. In addition, these results begin to define and measure a component of the NESF that had not been previously examined; the role of the teacher. The findings of this study support previous research related to the design features and outcomes of the NESF.

Replication of this study in order to gain more substantial evidence in support the results specifically related to the degree of importance of particular teaching behaviors/strategies is needed. Although the finding of the importance of feedback and debriefing to the participants of this study confirmed results from other simulation studies, the dimensionality of the SPECTS and the ranking of other teaching behaviors/strategies need further study.

A future quasi-experimental research design using the SPETCS could provide a means to test different simulation types and designs which will be useful as simulations as a teaching/learning method continues to grow in popularity in nursing education. Additionally, the examination of the relationship between effective teaching and the attainment of program outcomes is another area for additional study. Based on the tenets of learning theory and the NSEF, it is hypothesized that better teaching should produce better outcomes. Correlational studies may be developed to compare teaching and outcomes potentially across the different types of nursing programs.

In conclusion, there exists a strong initiative from the nursing education community to incorporate evidence-based educational practices into the curriculum and this study provides evidence that informs the role of the teacher and can serve as a basis to define best practices in teaching in simulation. This study applied current measurement theory with a strong theoretical foundation to develop an instrument to assess effective teaching in clinical simulation. The result was an easy to administer, 33 item instrument with robust, early evidence of reliability and validity which supports the professional development of faculty teaching in clinical simulation environments.

APPENDIX A

Student Perception of Effective Teaching in Clinical Simulation Scale

Student Perception of Effective Teaching in Clinical Simulation Scale

Directions: Using the 5 point scales below circle the numbers or letters that reflect your agreement or disagreement with each item and how important each item is for meeting the learning objectives of this simulation.

Extent of agreement:	Importance:
SD – strongly disagree D – disagree N – neutral (neither agree or disagree) A – agree SA – strongly agree	1 – not important 2 – slightly important 3 – moderately important 4 – very important 5 – extremely important

	Extent of agreement					Importance				
1. The instructor helped too much during the simulation.	SD	D	N	A	SA	1	2	3	4	5
2. The instructor allowed me time to think through challenging areas of the simulation.	SD	D	N	A	SA	1	2	3	4	5
3. Questions asked by the instructor after the simulation helped guide my thinking about the simulation experience.	SD	D	N	A	SA	1	2	3	4	5
4. The instructor provides me enough autonomy in the simulation to promote my learning.	SD	D	N	A	SA	1	2	3	4	5
5. The instructor provided useful feedback after the simulation.	SD	D	N	A	SA	1	2	3	4	5
6. The instructor facilitated my learning in this simulation.	SD	D	N	A	SA	1	2	3	4	5
7. Discussing the simulation during debriefing supports my understanding and reasoning.	SD	D	N	A	SA	1	2	3	4	5
8. Cues (hints) provided during the simulation guided my thinking.	SD	D	N	A	SA	1	2	3	4	5
9. An instructor-led debriefing is an important aspect of my simulation experience.	SD	D	N	A	SA	1	2	3	4	5
10. The instructor was comfortable with the simulation experience.	SD	D	N	A	SA	1	2	3	4	5
11. The simulation was interesting.	SD	D	N	A	SA	1	2	3	4	5

	Extent of Agreement					Importance				
12. Appropriate questions were asked during the debriefing of the simulation experience	SD	D	N	A	SA	1	2	3	4	5
13. The simulation was realistic.	SD	D	N	A	SA	1	2	3	4	5
14. I understood the objectives of the simulation.	SD	D	N	A	SA	1	2	3	4	5
15. The simulation fit with the objectives of this course.	SD	D	N	A	SA	1	2	3	4	5
16. I will be better able to care for a patient with this type of problem in clinical because I participated in this simulation.	SD	D	N	A	SA	1	2	3	4	5
17. Questioning by the instructor helps me to better understand the clinical situation experienced even though it is a simulated environment.	SD	D	N	A	SA	1	2	3	4	5
18. This simulation helped develop my critical thinking skills.	SD	D	N	A	SA	1	2	3	4	5
19. Cues were used in the simulation to help me progress through the experience.	SD	D	N	A	SA	1	2	3	4	5
20. The instructor served as a role model during the simulation.	SD	D	N	A	SA	1	2	3	4	5
21. The instructor demonstrated clinical expertise during this simulation experience.	SD	D	N	A	SA	1	2	3	4	5
22. The instructor was receptive to feedback.	SD	D	N	A	SA	1	2	3	4	5
23. Participation in this simulation was a valuable learning activity.	SD	D	N	A	SA	1	2	3	4	5
24. The instructor encouraged helpful collaboration among participants during debriefing.	SD	D	N	A	SA	1	2	3	4	5

	Extent of Agreement					Importance				
25. The difficulty of the simulation was appropriate.	SD	D	N	A	SA	1	2	3	4	5
26. Participation in clinical simulations helps me to meet clinical expectations when caring for real patients.	SD	D	N	A	SA	1	2	3	4	5
27. Cues were provided at appropriate times during the simulation.	SD	D	N	A	SA	1	2	3	4	5
28. Participation in this simulation helped me to understand classroom theory.	SD	D	N	A	SA	1	2	3	4	5
29. The instructor encouraged helpful collaboration among simulation participants during the simulation.	SD	D	N	A	SA	1	2	3	4	5
30. Clinical simulations are an effective learning strategy for me to problem-solve and to make decisions.	SD	D	N	A	SA	1	2	3	4	5
31. The clinical simulation flowed smoothly.	SD	D	N	A	SA	1	2	3	4	5
32. The instructor used a variety of questions during the debriefing	SD	D	N	A	SA	1	2	3	4	5
33. The clinical simulation experience was well-organized.	SD	D	N	A	SA	1	2	3	4	5
34. The instructor was enthusiastic during the simulation	SD	D	N	A	SA	1	2	3	4	5
35. Cues provided during the simulation supported my understanding.	SD	D	N	A	SA	1	2	3	4	5
36. My learning expectations were met in this clinical simulation	SD	D	N	A	SA	1	2	3	4	5
37. The simulation experience allows me to model a professional role in a realistic manner	SD	D	N	A	SA	1	2	3	4	5
38. Questions asked after the simulation helped me to understand the clinical decision-making necessary for this experience.	SD	D	N	A	SA	1	2	3	4	5

APPENDIX B

Recruitment letter to content experts

Dear Reviewer:

I am developing an instrument to measure effective teaching in clinical simulation within the context of undergraduate nursing education. The use of clinical simulations as a teaching/learning strategy in nursing education has grown dramatically in recent years, and a theory-based simulation model has been developed to guide the design, implementation, and evaluation of clinical simulations (Jeffries, 2005). However, the role of the teacher and teaching effectiveness are major components of the simulation model that have not been described or tested empirically. Many instruments are available to measure teaching effectiveness in classroom and clinical settings; however no instrument exists to measure this concept within simulated learning environments. Therefore, this scale has been developed in an effort to measure effective teaching in clinical simulations.

You are being asked to serve as a concept expert due to your recognition by the National League for Nursing as an expert in simulation. This designation supports your experience and expertise with clinical simulation. I am asking that you review and critique this instrument. Your contribution to the development of this instrument will serve to support the validity of a tool, which once created can serve as a tool for assessment, evaluation, and feedback in the ongoing professional development of nursing educators who use clinical simulations in the teaching-learning process.

Please refer to the attached form for your evaluation of the instrument. The conceptual definition and constructs used to develop the tool are also included as a guide when reviewing items. Please evaluate the representativeness and clarity of each statement in measuring the attributes of the concept as well as the comprehensiveness of the overall instrument to represent the total content domain. Space is included at the end of the form for your comments regarding the tool and suggestions for revisions.

I appreciate your time and effort providing feedback in the development of this instrument.

Sincerely,

Cynthia Reese, PhD (C), RN, CNE
Doctoral Student
Indiana University School of Nursing

The proposed instrument contains items based on constructs related to effective teaching identified in previous research and the literature. The instrument will have two response scales, Extent and Importance. The extent response scale relates to the student's perception of the extent to which the educator used the teaching strategy/behavior identified in each item. The importance response scale relates to student perceptions of the importance of the strategy identified in the item toward meeting the learning outcomes specified for the simulation experience. Items in both response scales will be measured using a five- point Likert scale. The extent response scale scores will range from a score of one indicating the strategy was not used at all to a score of five indicating the strategy was used to a great extent. The importance subscale will assess the importance of each strategy to achieve the learning objectives of the simulation as perceived by participants; a score of one indicates not at all important to a score of five indicating the strategy was extremely important.

Please use the following form to guide your review of items. Please consider:

1. Whether each item represents the concept domain.
2. If the content domain adequately address all dimensions of effective teaching.
3. If there are any style changes necessary in the wording of items.
4. If additional items are needed to improve the comprehensiveness of the instrument in representing the total content domain.
5. Do the items comprehensively represent the total content domain?

At the end of the grid there is space for additional comments.

Conceptual Definition of Effective Teaching in Clinical Simulation:

Effective teaching in clinical simulation is the degree to which the teaching strategies and behavioral characteristics of the instructor promote student achievement of the learning outcomes specified in the simulation experience.

Comments on conceptual definition of effective teaching:

APPENDIX C

Content Validity Grid

Content Validity Grid **Student Perception of Effective Teaching in Clinical Simulation Scale**					
Conceptual Definition: ***Effective teaching in clinical simulation is the degree to which the teaching strategies and behavioral characteristics of the instructor promote student achievement of the learning outcomes specified in the simulation experience.*** **Item**	**Representativeness** **1** = The item is **not representative** teaching effectiveness. **2** = The item **needs major revisions** to be representative of teaching effectiveness. **3** = The item **needs minor revisions** to be representative of teaching effectiveness. **4** = The item **is representative** of teaching effectiveness. **Please include your comments on items below.Space is provided at the end of the grid for suggestions for additions to this scale.**				**Please write the number of the related construct reflected in each item.** **1-Facilitator/ learner centered** **2-Feedback/ Debriefing** **3-Teaching Ability** **4-Modeling** **5-Interpersonal relationships** **6-Expectations** **7-Organization** **8-Cueing/ Student Support** **9-Questioning** **10-Enthusiasm**
I understood the objectives of the simulation.	**1**	**2**	**3**	**4**	
The instructor demonstrated clinical expertise.	**1**	**2**	**3**	**4**	
The instructor was fair.	**1**	**2**	**3**	**4**	
The instructor encouraged collaboration among simulation participants during the simulation.	**1**	**2**	**3**	**4**	
The difficulty of the simulation was appropriate.	**1**	**2**	**3**	**4**	
The instructor was organized during the clinical simulation.	**1**	**2**	**3**	**4**	
Cues (hints) were used in the simulation to help me progress through the experience.	**1**	**2**	**3**	**4**	

Questions asked by the instructor after the simulation helped guide my thinking.	**1**	**2**	**3**	**4**	
Participating in a simulation is a valuable teaching-learning activity for me.	**1**	**2**	**3**	**4**	
My learning expectations were met in the clinical simulation I experienced.	**1**	**2**	**3**	**4**	
The simulation was too difficult.	**1**	**2**	**3**	**4**	
The clinical simulation experience was well-organized.	**1**	**2**	**3**	**4**	
Cues provided during the simulation guided my thinking.	**1**	**2**	**3**	**4**	
The instructor allowed me time to think through challenging areas of the simulation.	**1**	**2**	**3**	**4**	
My instructor provides me enough autonomy in the simulation to promote my learning.	**1**	**2**	**3**	**4**	
My instructor serves as a professional role model for me when conducting the simulation.	**1**	**2**	**3**	**4**	
This simulation helped develop my critical thinking skills.	**1**	**2**	**3**	**4**	
The instructor was open-minded.	**1**	**2**	**3**	**4**	
The simulation was too easy.	**1**	**2**	**3**	**4**	
The instructor was organized during the clinical simulation.	**1**	**2**	**3**	**4**	
Clinical simulations are an effective learning strategy for me to problem-solve and to make decisions.	**1**	**2**	**3**	**4**	

Questioning by the instructor helps me to better understand the clinical situation experienced even though it is a simulated environment.	**1**	**2**	**3**	**4**	
Cues were provided at appropriate times during the simulation.	**1**	**2**	**3**	**4**	
The clinical simulation flowed smoothly.	**1**	**2**	**3**	**4**	
Participating in a simulation is a valuable teaching-learning activity for me.	**1**	**2**	**3**	**4**	
The simulation was too difficult.	**1**	**2**	**3**	**4**	
My learning expectations were met in the clinical simulation I experienced.	**1**	**2**	**3**	**4**	
The instructor encouraged collaboration among participants during debriefing.	**1**	**2**	**3**	**4**	
The instructor respected me during the simulation.	**1**	**2**	**3**	**4**	
The instructor was receptive to feedback.	**1**	**2**	**3**	**4**	
The simulation experience allows me to model a professional role in a realistic manner.	**1**	**2**	**3**	**4**	
The instructor was fair.	**1**	**2**	**3**	**4**	
The instructor served as a role model during the simulation.	**1**	**2**	**3**	**4**	
I will be better able to care for a patient with this type of problem in clinical because I participated in this simulation.	**1**	**2**	**3**	**4**	
The simulation was realistic.	**1**	**2**	**3**	**4**	
An instructor-led debriefing is an important aspect of my simulation experience.	**1**	**2**	**3**	**4**	
The instructor provided feedback after the simulation.	**1**	**2**	**3**	**4**	
The instructor facilitated my learning in this simulation.	**1**	**2**	**3**	**4**	
The instructor provided feedback during the simulation.	**1**	**2**	**3**	**4**	

The instructor helped me during times I felt challenged during this simulation.	**1**	**2**	**3**	**4**	
The instructor helped too much during the simulation.	**1**	**2**	**3**	**4**	
The instructor was comfortable using the simulator.	**1**	**2**	**3**	**4**	
The instructor demonstrated clinical expertise.	**1**	**2**	**3**	**4**	
The instructor respected the group during the simulation.	**1**	**2**	**3**	**4**	
The difficulty of the simulation was appropriate.	**1**	**2**	**3**	**4**	
The simulation was interesting.	**1**	**2**	**3**	**4**	
The clinical simulation experience was well-organized.	**1**	**2**	**3**	**4**	
My learning experience was organized and met my expectations.	**1**	**2**	**3**	**4**	
Cues (hints) used in the simulation helped me progress through the simulation.	**1**	**2**	**3**	**4**	
The instructor was enthusiastic during the simulation.	**1**	**2**	**3**	**4**	
Questions asked by the instructor after the simulation helped guide my thinking.	**1**	**2**	**3**	**4**	
Appropriate questions were asked during the debriefing of the simulation experience.	**1**	**2**	**3**	**4**	
Cues provide during the simulation supported my understanding.	**1**	**2**	**3**	**4**	
Participation in this simulation helped me to understand classroom theory.	**1**	**2**	**3**	**4**	
Participation in clinical simulations helped me to meet the clinical expectations I have when caring for real patients.	**1**	**2**	**3**	**4**	
There was strong student/faculty interaction during my clinical simulation experience.	**1**	**2**	**3**	**4**	

The instructor encouraged collaboration among simulation participants during the simulation.	**1**	**2**	**3**	**4**	
The instructor was flexible.	**1**	**2**	**3**	**4**	
The simulation was interesting.	**1**	**2**	**3**	**4**	
The simulation fit with the objectives of this course.	**1**	**2**	**3**	**4**	
The simulation was interesting.	**1**	**2**	**3**	**4**	
Discussing the simulation during debriefing supports my understanding and reasoning.	**1**	**2**	**3**	**4**	
The feedback from the instructor during the simulation helped my understanding of the simulation.	**1**	**2**	**3**	**4**	
I felt supported (helped) by the instructor during this simulation.	**1**	**2**	**3**	**4**	
I understood the objectives of the simulation.	**1**	**2**	**3**	**4**	
Cues provided during the simulation guided my thinking.	**1**	**2**	**3**	**4**	
Questions asked after the simulation helped me to understand the required clinical decision-making for this experience.	**1**	**2**	**3**	**4**	
The instructor used a variety of questions during the debriefing.	**1**	**2**	**3**	**4**	
The instructor was enthusiastic and dynamic.	**1**	**2**	**3**	**4**	
The instructor moved in too quickly when I was having difficulty during the simulation.	**1**	**2**	**3**	**4**	

(Adapted from Grant & Davis, 1997)

Please include additional comments here:

____

____

____

____

____

APPENDIX D

Permission to use instruments

Vancouver, Jan 2nd 2008
Cynthia Reese,MS, RN, CNE
Professor of Nursing
Lincoln Land Community College
5250 Shepherd Rd
PO Box19256
Springfield, Ill
USA62794-9256

Dear Ms. Reese,

You have my permission to use either version of the tool **Nursing Clinical Teacher Effectiveness Inventory** My co-investigator agreed that I might allow the use of the tool for anyone I feel is using it in an ethical fashion. Since you are conducting the research within a University setting, I assume that you have ethical approval from the Universities ethics committee.

Yours sincerely

Judith Mogan, RN. MA (Ad Ed)
Assistant Professor Emerita

APPENDIX E

Nursing Clinical Teaching Effectiveness Inventory

Nursing Clinical Teaching Effectiveness Inventory

Best Clinical Teacher

DIRECTIONS Picture the best clinical teacher you have ever had. Think back specifically what this person did to make him/her the best clinical teacher. For each statement circle the number which indicates how descriptive the practice is of this individual

Teaching Behaviors		Not at All			Descriptive			Very
Teaching Ability								
1.	Explains clearly	1	2	3	4	5	6	7
2.	Emphasizes what is important	1	2	3	4	5	6	7
3.	Stimulates student interest in the subject	1	2	3	4	5	6	7
4.	Remains accessible to students	1	2	3	4	5	6	7
5.	Demonstrates clinical procedures and techniques	1	2	3	4	5	6	7
6.	Guides students' development of clinical skills	1	2	3	4	5	6	7
7.	Provides specific practice opportunity	1	2	3	4	5	6	7
8.	Offers special help when difficulties arise	1	2	3	4	5	6	7
9.	Is well prepared for teaching	1	2	3	4	5	6	7
10.	Enjoys teaching	1	2	3	4	5	6	7
11.	Encourages active participation in discussion	1	2	3	4	5	6	7
12.	Gears instruction to students level of readiness	1	2	3	4	5	6	7
13.	Quickly grasps what students are asking or telling	1	2	3	4	5	6	7
14.	Answers carefully and precisely questions raised by students	1	2	3	4	5	6	7
15.	Questions students to elicit underlying reasoning	1	2	3	4	5	6	7
16.	Helps students organize their thoughts about patient problems	1	2	3	4	5	6	7
17.	Promotes student independence	1	2	3	4	5	6	7
Nursing Competence		1	2	3	4	5	6	7
18.	Demonstrates clinical skill and judgment	1	2	3	4	5	6	7
19.	Demonstrates communication skills	1	2	3	4	5	6	7
20.	Reveals broad reading in his/her area of interest	1	2	3	4	5	6	7
21.	Discusses current development in his/her field	1	2	3	4	5	6	7
22.	Directs students to useful literature in nursing	1	2	3	4	5	6	7
23.	Demonstrates a breadth of knowledge in nursing	1	2	3	4	5	6	7
24.	Recognizes own limitations	1	2	3	4	5	6	7
25.	Takes responsibility of own actions	1	2	3	4	5	6	7

26.	Is a good role model	1	2	3	4	5	6	7
<u>Evaluation</u>		1	2	3	4	5	6	7
27.	Makes specific suggestions for improvement	1	2	3	4	5	6	7
28.	Provides frequent feedback on students' performance	1	2	3	4	5	6	7
29.	Identifies students' strengths and limitations objectively	1	2	3	4	5	6	7
30.	Observes students' performance frequently	1	2	3	4	5	6	7
31.	Communicates expectations of students	1	2	3	4	5	6	7
32.	Gives students positive reinforcement for good contributions, observations or performance	1	2	3	4	5	6	7
33.	Corrects students' mistakes without belittling them	1	2	3	4	5	6	7
34.	Does not criticize students in front of others	1	2	3	4	5	6	7
<u>Interpersonal Relations</u>		1	2	3	4	5	6	7
35.	Provides support and encouragement to students	1	2	3	4	5	6	7
36.	Is approachable	1	2	3	4	5	6	7
37.	Encourages a climate of mutual respect	1	2	3	4	5	6	7
38.	Listens attentively	1	2	3	4	5	6	7
39.	Shows a personal interest in students	1	2	3	4	5	6	7
40.	Demonstrates empathy	1	2	3	4	5	6	7
<u>Personality</u>		1	2	3	4	5	6	7
41.	Demonstrates enthusiasm	1	2	3	4	5	6	7
42.	Is a dynamic and energetic person	1	2	3	4	5	6	7
43.	Self –confidence	1	2	3	4	5	6	7
44.	Is self-critical	1	2	3	4	5	6	7
45.	Is open-minded and non-judgemental	1	2	3	4	5	6	7
46.	Has a good sense of humour	1	2	3	4	5	6	7
47.	Appeasers organized	1	2	3	4	5	6	7

APPENDIX F

Student's Evaluation of Educational Quality

Student's Evaluation of Educational Quality

Using the 5 point scale below, indicate by circling the most appropriate number the extent of your agreement or disagreement with the factors/statements listed below. Try to relate your answers to the current simulation as much as possible.

LEARNING	**Strongly Disagree**	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
1. You find the course intellectually challenging and stimulating.	1	2	3	4	5
2. You have learned something which you consider valuable.	1	2	3	4	5
3. Your interest in the subject increased as a consequence of this course.	1	2	3	4	5
4. You have learned and understood the subject materials of this course.	1	2	3	4	5
INDIVIDUAL RAPPORT					
5. Instructor is friendly towards individual students.	1	2	3	4	5
6. Instructor makes students feel welcome in seeking help/advice in or outside of class.	1	2	3	4	5
7. Instructor has a genuine interest in individual students.	1	2	3	4	5
8. Instructor is adequately accessible to students during office hours or after class.	1	2	3	4	5
ENTHUSIASM					
9. Instructor is enthusiastic about teaching this course.	1	2	3	4	5
10. Instructor is dynamic and energetic in conducting the course.	1	2	3	4	5
11. Instructor enhances presentations with the use of humor.	1	2	3	4	5
12. Instructor's style of presentation held my students' interest during class.	1	2	3	4	5
EXAMINATIONS					
13. Feedback on examinations / graded materials is valuable.	1	2	3	4	5
14. Methods of evaluating student work are fair and appropriate.	1	2	3	4	5
15. Examinations/graded materials tested course content as emphasized by instructor.	1	2	3	4	5
ORGANIZATION					
16. Instructor's explanations are clear.	1	2	3	4	5
17. Course materials are well prepared and carefully explained.	1	2	3	4	5
18. Proposed objectives agree with those actually taught so you know where the course is going.	1	2	3	4	5
19. Instructor gives presentations that facilitate taking notes.	1	2	3	4	5
BREADTH					
20. Instructor contrasts the implications of various theories.	1	2	3	4	5
21. Instructor presents the background or origin of ideas/concepts developed in class.	1	2	3	4	5
22. Instructor presents points of view other than his/her own when appropriate.	1	2	3	4	5
23. Instructor adequately discusses current developments in the field.	1	2	3	4	5

GROUP INTERACTION

24. Students are encouraged to participate in class discussions.	1	2	3	4	5
25. Students are invited to share their ideas and knowledge.	1	2	3	4	5
26. Students are encouraged to ask questions and were given meaningful answers.	1	2	3	4	5
27. Students are encouraged to express their own ideas and/or question the Instructor.	1	2	3	4	5

ASSIGNMENTS

28. Required readings / texts are valuable.	1	2	3	4	5
29. Readings, homework, laboratories contribute to appreciation and understanding of the subject.	1	2	3	4	5

IMPORTANT COMPONENTS OF TEACHING EFFECTIVENESS
Please indicate how important you consider the following factors to be in teaching this course effectively

(1 = not important, 2=somewhat important, 3=moderately important, 4=very important, 5=extremely important)

Stimulate learning / Academic Value	1	2	3	4	5
Examinations / Grading	1	2	3	4	5
Group Interaction	1	2	3	4	5
Instructor Enthusiasm	1	2	3	4	5
Organization / Clarity	1	2	3	4	5
Assignments / Reading	1	2	3	4	5
Individual rapport with students	1	2	3	4	5
Breadth of coverage	1	2	3	4	5
Appropriate workload / difficulty	1	2	3	4	5

APPENDIX G

33 item Student Perception of Effective Teaching in Clinical Simulation Scale

Student Perception of Effective Teaching in Clinical Simulation Scale

Directions: Using the 5 point scales below circle the numbers or letters that reflect your agreement or disagreement with each item and how important each item is for meeting the learning objectives of this simulation.

Extent of agreement: SD – strongly disagree D – disagree N – neutral (neither agree or disagree) A – agree SA – strongly agree		**Importance:** 1 – not important 2 – slightly important 3 – moderately important 4 – very important 5 – extremely important
	Extent of Agreement	**Importance**
1. The instructor allowed me time to think through challenging areas of the simulation.	SD D N A SA	1 2 3 4 5
2. Questions asked by the instructor after the simulation helped guide my thinking about the simulation experience.	SD D N A SA	1 2 3 4 5
3. The instructor provides me enough autonomy in the simulation to promote my learning.	SD D N A SA	1 2 3 4 5
4. The instructor provided useful feedback after the simulation.	SD D N A SA	1 2 3 4 5
5. The instructor facilitated my learning in this simulation.	SD D N A SA	1 2 3 4 5
6. Discussing the simulation during debriefing supports my understanding and reasoning.	SD D N A SA	1 2 3 4 5
7. An instructor-led debriefing is an important aspect of my simulation experience.	SD D N A SA	1 2 3 4 5
8. The instructor was comfortable with the simulation experience.	SD D N A SA	1 2 3 4 5
9. The simulation was interesting.	SD D N A SA	1 2 3 4 5
10. Appropriate questions were asked during the debriefing of the simulation experience	SD D N A SA	1 2 3 4 5

	Extent of Agreement	Importance
11. The simulation was realistic.	**SD D N A SA**	**1 2 3 4 5**
12. The simulation fit with the objectives of this course.	**SD D N A SA**	**1 2 3 4 5**
13. I will be better able to care for a patient with this type of problem in clinical because I participated in this simulation.	**SD D N A SA**	**1 2 3 4 5**
14. Questioning by the instructor helps me to better understand the clinical situation experienced even though it is a simulated environment.	**SD D N A SA**	**1 2 3 4 5**
15. This simulation helped develop my critical thinking skills.	**SD D N A SA**	**1 2 3 4 5**
16. Cues were used in the simulation to help me progress through the experience.	**SD D N A SA**	**1 2 3 4 5**
17. The instructor served as a role model during the simulation.	**SD D N A SA**	**1 2 3 4 5**
18. The instructor demonstrated clinical expertise during this simulation experience.	**SD D N A SA**	**1 2 3 4 5**
19. The instructor was receptive to feedback.	**SD D N A SA**	**1 2 3 4 5**
20. Participation in this simulation was a valuable learning activity.	**SD D N A SA**	**1 2 3 4 5**
21. The instructor encouraged helpful collaboration among participants during debriefing.	**SD D N A SA**	**1 2 3 4 5**
22. The difficulty of the simulation was appropriate.	**SD D N A SA**	**1 2 3 4 5**
23. Participation in clinical simulations helps me to meet clinical expectations when caring for real patients.	**SD D N A SA**	**1 2 3 4 5**

	Extent of Agreement	Importance
24. Cues were provided at appropriate times during the simulation.	**SD D N A SA**	**1 2 3 4 5**
25. Participation in this simulation helped me to understand classroom theory.	**SD D N A SA**	**1 2 3 4 5**
26. The instructor encouraged helpful collaboration among simulation participants during the simulation.	**SD D N A SA**	**1 2 3 4 5**
27. Clinical simulations are an effective learning strategy for me to problem-solve and to make decisions.	**SD D N A SA**	**1 2 3 4 5**
28. The instructor used a variety of questions during the debriefing	**SD D N A SA**	**1 2 3 4 5**
29. The clinical simulation experience was well-organized.	**SD D N A SA**	**1 2 3 4 5**
30. The instructor was enthusiastic during the simulation.	**SD D N A SA**	**1 2 3 4 5**
31. My learning expectations were met in this clinical simulation	**SD D N A SA**	**1 2 3 4 5**
32. The simulation experience allows me to model a professional role in a realistic manner	**SD D N A SA**	**1 2 3 4 5**
33. Questions asked after the simulation helped me to understand the clinical decision-making necessary for this experience.	**SD D N A SA**	**1 2 3 4 5**

APPENDIX H

Institutional Review Board Approval

INTERDEPARTMENTAL COMMUNICATION

Research Compliance Administration
Indiana University-Purdue University Indianapolis

DATE:	February 24, 2008
TO:	Dr. Pamela Jeffries Nursing NU 140 IUPUI
FROM:	Regina Wininger Research Compliance Administration UN 618 IUPUI
SUBJECT:	IUPUI/Clarian Institutional Review Committee - Exempt Review of Human Study
Study No.:	EX0802-53B
Study Title:	"Effective Teaching in Clinical Simulation: Development of the Student's Perception of Effective Teaching in Clinical Simulations Scale"

Your application for approval of the study named above has been accepted as meeting the criteria of exempt research as described by Federal Regulations [45 CFR 46.101(b), paragraph 1]. A copy of the acceptance is enclosed for your file. If the research is conducted at or funded by the VA, research may not be initiated until approval is received from the VA Research and Development Committee.

Please contact the Indiana University School of Medicine Office of Compliance Services for information regarding a Data Use Agreement, if applicable.

Although a continuing review is not required for an exempt study, prior approval must be obtained before change(s) to the originally approved study can be initiated. When you have completed your study, please inform our office in writing.

When corresponding with our office regarding this study, please refer to the exact study number and title.

If you should have any questions, please contact our office at 317-274-8289.

Enclosures: ☒ Copy of acceptance

APPENDIX I

Informed Consent Statement

IUPUI and CLARIAN INFORMED CONSENT STATEMENT FORM

Effective Teaching in Clinical Simulation: Development of the Student Perception of Effective Teaching in Clinical Simulation Scale.

You are invited to participate in a research study to develop an instrument to measure effective teaching in clinical simulations. You were selected as a possible subject because you are an undergraduate nursing student at IUSON who is participating in clinical simulations during your course of study**.** We ask that you read this form and ask any questions you may have before agreeing to be in the study.

The study is being conducted by Dr. Pamela Jeffries, Associate Dean for Undergraduate Programs, IUSON and Cynthia Reese, doctoral student at IUSON.

STUDY PURPOSE

The purpose of this study is to create a survey that measures effective teaching in clinical simulation contexts.

NUMBER OF PEOPLE TAKING PART IN THE STUDY:

If you agree to participate, you will be one of 100 nursing students who will be participating in this research.

PROCEDURES FOR THE STUDY:

If you agree to be in the study, you will do the following things:

You will complete 3 different surveys following participation in the clinical simulation. It is anticipated that it will take no longer than 30 minutes to complete all of the surveys. Approximately two weeks following the simulation, you will retake one of the surveys during a clinical post-conference. It should take no longer than 10 – 15 minutes for you to complete the survey.

RISKS OF TAKING PART IN THE STUDY:

There is very little risk to you as a participant in this teaching/learning experience and evaluation. Participation in the study is voluntary, but the simulation and debriefing are required components in the course. Therefore, all students will have this experience as part of their course to promote learning. It will be the student's option whether to complete the questionnaires for the study or not. The choice to complete the survey will have no impact on your course grade,you're your responses will be confidential.

BENEFITS OF TAKING PART IN THE STUDY:

The benefit to participation that is reasonable to expect is that your participation in this study will aid in the development of a tool to improve teaching strategies of nurse educators who use clinical simulations in the teaching-learning process.

ALTERNATIVES TO TAKING PART IN THE STUDY:

Instead of taking part in the study, you have the option not to participate.

CONFIDENTIALITY

Efforts will be made to keep your personal information confidential. We cannot guarantee absolute confidentiality. Your personal information may be disclosed if required by law. Your identity will be held in confidence in reports in which the study may be published and databases in which results may be stored..

Organizations that may inspect and/or copy your research records for quality assurance and data analysis include groups such as the study investigator and his/her research associates, the IUPUI/Clarian Institutional Review Board or its designees.

PAYMENT

You will not receive payment for taking part in this study.

CONTACTS FOR QUESTIONS OR PROBLEMS

For questions about the study or a research-related injury, contact the researcher Cynthia Reese at 217-825-4734 or Dr. Pamela Jeffries at 317-274-2805. If you cannot reach the researcher during regular business hours (i.e. 8:00AM-5:00PM), please call the IUPUI/Clarian Research Compliance Administration office at (317) 278-3458 or (800) 696-2949. After business hours, please call Cynthia Reese at 217-825-4734.

In the event of an emergency, you may contact Cynthia Reese at 217-825-4734.

For questions about your rights as a research participant or to discuss problems, complaints or concerns about a research study, or to obtain information, or offer input, contact the IUPUI/Clarian Research Compliance Administration office at (317) 278-3458 or (800) 696-2949.

VOLUNTARY NATURE OF STUDY

Taking part in this study is voluntary. You may choose not to take part or may leave the study at any time. Leaving the study will not result in any penalty or loss of benefits to which you are entitled. Your decision whether or not to participate in this study will not affect your current or future relations with Indiana University School of Nursing.

SUBJECT'S CONSENT
In consideration of all of the above, I give my consent to participate in this research study.

I will be given a copy of this informed consent document to keep for my records. I agree to take part in this study.

Subject's Printed Name:____

Subject's Signature:________________________________ Date:________________

Print**ed Name of Person Obtaining Consent:**______________________________________

Signature of Person Obtaining Consent**:**____

Date**:**________________

APPENDIX J

Student and Instructor demographic surveys

Student Demographic Survey

Please complete the following:

Age: ______

Gender:

__________Female

__________Male

Ethnicity:

__________Hispanic or Latino

__________Not Hispanic or Latino

Race:

__________American Indian or Alaska Native

__________Asian

__________Black or African American

__________Native Hawaiian or other Pacific Islander

__________White

__________Other or Unknown: Please specify____________________

Current Grade Point Average:____

Other college degree:

__________Yes: Please specify degree held_________________________

__________No

Previous experience with clinical simulations:

__________Yes: If yes,

Please specify number of simulations you have participated in _______

__________No

Years of experience working in healthcare:

____.

Instructor Demographic Survey

Please complete the following:

Age: ________________

Gender:

___________Female

___________Male

Ethnicity:

___________Hispanic or Latino

___________Not Hispanic or Latino

Race:

___________American Indian or Alaska Native

___________Asian

___________Black or African American

___________Native Hawaiian or other Pacific Islander

___________White

___________Other or Unknown: Please specify______________________

Current professional certifications: ____.

Educational Preparation and year of graduation:____.

Years of teaching experience: ____.

Years of teaching experience with clinical simulations:___________________________________.

REFERENCES

Alfaro-LeFevre, R. (2004). *Critical thinking and clinical judgment.* St. Louis, MO: Saunders.

American Association of Colleges of Nursing (2005). *Faculty shortages in baccalaureate and graduate nursing programs: the scope of the problem and strategies for expanding the supply.* AACN White Paper. Washington D.C.

American Association of Medical Colleges (2000). 2010 scenarios: education. AAMC White Paper.

Aronson, B., Rosa, J., Anfinson, J. & Light, N. (1997). A simulated clinical problem solving experience. *Nurse Educator, 22*(6), 17-19.

Bain, K. (2004). *What the best college teachers do*. Cambridge, MA: Harvard University Press.

Bearnson, C. & Wilker, K. (2005). Human patient simulators: a new face in baccalaureate nursing education at Brigham Young University. *Journal of Nursing Education, 44*(9), 421-425.

Becker, K. L., Rose, L. E., Berg, J. B., Park, H. & Shatzer, J. H. (2006). The teaching effectiveness of standardized patients. *Journal of Nursing Education, 45*(4).

Beckman, T. J., Ghosh, A. K., Cook, D. A., Erwin, P. J. & Mandrekar, J. N. (2004). How reliable are assessments of clinical teaching? A review of published instruments. *Journal of General Internal Medicine, 19,* 971-977.

Beitz, J. M. & Wieland, D. (2005). Analyzing the teaching effectiveness of clinical nursing faculty of full- and part-time generic BSN, LPN-BSN, and RN-BSN nursing students. *Journal of Professional Nursing, 21*(1).

Bienstock, J. L., Katz, N. T., Cox, S. M., Hueppchen, N., Erickson, S., Pushcheck, E. E. (2007). To the point: medical education reviews-providing feedback. *American Journal of Obstetrics & Gynecology, 196*(6), 508-513.

Billings, D. & Halstead, J. eds. (2004) *Teaching in nursing: a guide for faculty* (2nd ed.). Philadelphia: W.B. Saunders.

Bloom, B.S. (Ed) (1957). *Taxonomy of educational objectives. The classification of educational goals.* New York: McKay.

Bonk, C. J. & Cunningham, D. J. (1998). *Searching for learner-centered, constructivist, and sociocultural components of collaborative education learning tools in Electronic Collaborators*. Mahwah, NJ: Lawrence Erlhbaum Association, Inc.

Bradley, P. (2006). The history of simulation in medical education and possible future directions. *Medical Education, 40,* 254-262.

Bremner, M. N., Auddell, K., Bennett, D. N. & VanGeest, J. B. (2006). The use of human patient simulators: best practices with novice nursing students. *Nurse Educator, 31*(4).

Brown, S. T. (1981). Faculty and student perceptions of effective clinical instruction. *Journal of Nursing Education, 20*(9), 4-15.

Buskist, W. (2002). Effective teaching: perspectives and insights from division two's 2 and 4 year awardees. *Teaching of psychology, 29*(3).

Capra, F. (2002). *The hidden connections: integrating the biological, cognitive, and social dimensions of life into a science of sustainability*. New York: Doubleday.

Carmines, E. G. & Zeller, R. A. (1979). *Reliability and validity assessment*. Thousand Oaks: Sage Publications.

Chau, J., Chang, A., Lee, I., Ip, W., Lee, D. & Wooton, Y. (2001). Effects of using videotaped vignettes on enhancing students' critical thinking ability in a baccalaureate nursing programme. *Journal of Advanced Nursing, 36*, 112-119.

Chickering, A. W. & Gamson, Z. F. (1987). Seven principles of good practice in undergraduate education. *AAHE Bulletin, 39*(7), 5-10.

Chickering, A. & Ehrmann, S. C. (1996). Implementing the seven principles: Technology as a lever. *AAHE Bulletin* (October), 3-6.

Chickering, A. W. & Gamson, Z. F. (1999). Development and adaptations of the seven principles of good practice in undergraduate education. *New Directions for Teaching and Learning, 80*(Winter), 75-81.

Cioffi, J. (2001). Clinical simulations: development and validation. *Nurse Education Today, 21*, 477-486.

Cioffi, J., Purcal, N. & Arundell, F. (2005). A pilot study to investigate the effect of a simulation strategy on the clinical decision making of midwifery students. *Journal of Nursing Education, 44*(3), 131-134.

Comer, S. (2005). Patient care simulations: role playing to increase understanding. *Nursing Educational Perspectives, 26*(6).

Copeland, H. L. & Hewson, M. G. (2000). Developing and testing an instrument to measure the effectiveness of clinical teaching at an academic medical center. *Academic Medicine, 75*(2), 161-166.

D'Apollonia, S. & Abrami, P.C, (1997). Navigating student ratings of instruction. *American Psychologist, 52,*1198-1208.

DeVellis, R. F. (2003). *Scale Development: Theory and Applications* (Second Edition ed.). Thousand Oaks: Sage Publications.

DeVon, H., Block, M., Moyle-Wright, P., Ernst, D., Hayden, S., Lazzara, D., Savoy, S., Kostas-Polston, E. (2007). Clinical scholarship: A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship, 39(*2), 155-164.

Dillman, D. A., (2000). *Mail and internet* surveys (2nd ed.). New York: Wiley.

Dunphy, B. C. & Dunphy, S. L., (2003). Assisted performance and the zone of proximal development: a potential framework for providing surgical education. *Australian Journal of Educational and Developmental Psychology, 3*, 48-58

Feingold, C., Calaluce, M. & Kallen, M. (2004). Computerized patient model and simulated clinical experiences: evaluation with baccalaureate nursing students. *Journal of Nursing Education, 43*(4), 156 163.

Ferketich, S. (1991). Aspects of item analysis. *Research in Nursing and Health, 14*, 165-168.

Field, A. (2005). *Discovering Statistics Using SPSS* (2nd ed). Thousand Oaks: Sage Publications.

Fong, C. M. & McCauley, G. T. (1993). Measuring the nursing, teaching, and interpersonal effectiveness of clinical instructors. *Journal of Nursing Education, 32*(7).

Fosnot, C. T. (2005). *Constructivism: theory, perspectives, and practice.* New York: Teacher's College Press.

Grant, J. S. & Davis, L. L. (1997). Focus on quantitative methods: Selection and use of content experts for instrument development. *Research in Nursing and Health, 20,* 269-274.

Greenwood J. (2000). Critical thinking and nursing scripts: the case for the development of both. *Journal of Advanced Nursing, 31*(2), 428-436.

Greiner, A. C. & Knebel, E., (Eds.) (2003). *Health Professions Education: a bridge to quality*.Washington, D.C.: Institute of Medicine National Academies Press.

Hertel, J. P. & Mills, B. J. (2002). *Using simulations to promote learning in higher education: an introduction*. Sterling, VA: Stylus Publishing.

Hotchkiss, M., Biddle, C. & Fallacaro, M. (2002). Assessing the authenticity of the human simulation experience in anesthesiology. *AANA Journal, 70*(6), 470-473.

Jeffries, P. (2005). A framework for designing implementing and evaluating simulations used as teaching strategies in nursing. *Nursing Education Perspectives, 26*(2), 96 103.

Jeffries, P. (2005). Technology trends in nursing education: next steps. *Journal of Nursing Education, 44*(1), 3-4.

Jeffries, P. R. & Rizzolo, M. A. (2006). *Designing and implementing models for the innovative use of simulation to teach nursing care of ill adults and children: a national, multi-site, multi-method study*. New York: National League for Nursing.

Jeffries, P. R. (Ed.). (2007). *Simulation in nursing education from conceptualization to evaluation*. New York: National League for Nursing.

Johnson, J., Zerwic, J. & Theis, S. (1999). Clinical simulation laboratory: an adjunct to clinical teaching. *Nurse Educator, 24*(5), 37-41.

Keeley, J., Smith, D. & Buskist, W. (2006). The teacher behaviors checklist: factor analysis of its utility for evaluating teaching. *Teaching of Psychology*, *33*(2), 84-91.

Kerlinger, F. N. & Lee, H. B. (2000). *Foundations of Behavioral Research* (4th ed.). Orlando: Harcourt College Publishers.

Kirschling, J. M., Fields, J., Imle, M., Mowery, M., Tanner, C., Perrin, N., et al. (1995). Evaluating teaching effectiveness. *Journal of Nursing Education, 34*(9), 401-410.

Knox, J. E. & Mogan, J. (1985). Important clinical teacher behaviors as perceived by university nursing faculty, students and graduates. *Journal of Advanced Nursing, 10*, 25-30.

Kuh, G. D., Pace, C. & Vesper, N. (1997). The development of process indicators to estimate student gains associated with good practices in undergraduate education. *Research in Higher Education, 38*(4), 435-454.

Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Larew, C., Lessans, S., Spunt, D., Foster, D. & Covington, B. G. (2007). Innovations in clinical simulation: application of Benner's theory in an interactive patient care simulation. *Nursing Education Perspectives, 27*(1).

Lattuca, L. R. (2002). Learning interdisciplinarity: sociocultural perspectives on academic work. *The Journal of Higher Education, 73*(6), 711-739.

Lattuca, L. R. (2006). The constructivist pedagogy we're looking for. *Journalism and Mass Communication Educator, Winter 2006*.

Leamon M. H. & Fields L. (2005) Measuring teaching effectiveness in a pre-clinical multi-instructor course: a case study in the development and application of a Brief Instructor Rating Scale. *Teaching and Learning Medicine, 17*(2), 119-29.

Lonser, V., Abbott, R., Allen, K. & Davidhizar, R. (2006). Implementation of problem-based learning in a final semester comprehensive nursing course. *The Health Care Manager*, *25*(2), 184-193.

Lynn, M. (1986). Determination and quantification of content validity. *Nursing Research*, *35*, 382-385.

Marsh, H. W. (1987). Students' evaluations of university teaching: research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, *11*(3), 253-388.

Marsh, H.W. & Roche, L. (1997). Making students' evaluations of teaching effectiveness effective: the critical issues of validity, bias, and utility. *The American Psychologist*, *52*(11), 1187-1197.

McCausland, L., Curran, C. & Cataldi, P. (2004). Use of a human simulator for undergraduate nurse education. *International Journal of Nursing Education Scholarship*, *1*(1).

McKeachie, W. J.(1997). Student ratings: the validity of use. *American Psychologist*, *52*(11).

Medley, C. & Horne, C. (2005). Using simulation technology for undergraduate nursing education. *Journal of Nursing Education*, *44*(1), 31-34.

Mogan, J. & Knox, J. E. (1987). Characteristics of 'best' and 'worst' clinical teachers as percieved by university nursing faculty and students. *Journal of Advanced Nursing, 12,* 331-337.

Mourad, O. & Redelmeier, D. A. (2006). Clinical teaching and clinical outcomes: teaching capability and its association with patient outcomes. *Medical Education, 40*, 637-644.

Munro, B. (2005). *Statistical methods for healthcare research* (4th ed.). Philadelphia: Lippincott, Williams & Wilkins.

National League for Nursing. (2005). *Transforming nursing education.* [Position Statement]. New York: National League for Nursing.

Nehring, W., Lashley, F. & Ellis, W. (2002). Critical incident nursing management using human patient simulators. *Nursing Education Perspectives*, *23*(3), 128-132.

Nehring W.M. & Lashley, F.R.. (2004). Current use and opinions regarding human patient simulators in nursing education: an international survey. *Nursing Education Perspectives. 25*(5), 244-248.

Netemeyer, R. G., Bearden, W. O. & Sharma, S. (2003). *Scaling procedures: Issues and applications*. Thousand Oaks: Sage Publications.

Nunnally, J. C. (1978). *Psychometric Theory (2nd ed.).* New York: McGraw-Hill.

Organized Dictionary.com. *The American Heritage® Dictionary of the English Language*, *Fourth Edition*. Houghton Mifflin Company, 2004. http://dictionary.reference.com/browse/organized (accessed: January 8, 2008).

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Reviews of Psychology, 49,* 347-375.

Pesut, D. & Herman, J., (1999). *Clinical reasoning: the art and science of critical and creative thinking.* Albany, NY: Delmar Publishers.

Raingruber B. & Bowles, K. (2000). Developing student evaluation instruments to measure instructor effectiveness. *Nurse Educator, 25*(2), 65-69.

Rauen, C. (2001). Using simulation to teach critical thinking skills. *Critical Care Nursing Clinics of North America 13*(1), 93-103.

Ravert, P. (2002). An integtrative review of computer-based simulation in the education process. *Computers, Informatics, Nursing*, *20*(5), 203-208.

Reed, D., Price, E., Windish, D., Wright, S., Gozu, A., Hsu, E., et al. (2005). Challenges in systematic reviews of educational intervention studies. *Annals of Internal Medicine*, *142*, 1080-1089.

Rudolph, J., Simon, R., Dufresne, R. & Raemer, D. (2006). There's no such thing as nonjudgemental debriefing: a theory and method for debriefing with good judgment. *Simulation in Healthcare, 1*(1), 49-55.

Richardson, V. (2003). Constructivist Pedagogy. *Teachers college record, 105*(9), 1623-1640.

Sanders, D. & Welk, D. S. (2005). Strategies to scaffold student learning: applying Vygotsky's zone of proximal development. *Nurse Educator*, *30*(5), 203-207.

Schaeffer, G., Epting, K., Zinn, T. & Buskist, W. (2003). Student and faculty perceptions of effective teaching: a successful replication. *Teaching of Psychology*, *30*(2).

Scherer, Y., Bruce, S., Graves, B. & Erdley, W. S. (2003). Acute care nurse practitioner education: enhancing performance through the use of clinical simulation. *AACN Clinical Issues*, *14*(3), 331-341.

Schon, D. (1987). *Educating the reflective practitioner: toward a new design for teaching and learning in the professions.* San Francisco: Jossey-Bass.

Seropian, M. (2003). General concepts in full scale simulation: getting started. *Anesthesia Analog, 97,* 1695-1705.

Seropian, M., Brown, K., Gavilanes, J. & Driggers, B. (2004). Simulation: not just a manikin. *Journal of Nursing Education, 43*(4), 164-169.

Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428.

Simmons, B., Lanuza., Fonteyn, M., Hicks, F. & Holm, K. (2003). Clinical reasoning in experienced nurses. *Western Journal of Nursing Research, 25*(6), 701-724.

Sprunt, D., Foster, D. & Adams, K. (2004). Mock code: a clinical simulation module. *Nurse Educator, 29*(5), 192-194.

Streufert,S., Satish, U. & Barach, P. (2001). Improving medical care: the use of simulation technology. *Simulation & Gaming, 32*(2), 164-174.

Sudzina, M. R. (1997). Case study as a constructivist pedagogy for teaching educational psychology. *Educational Psychology Review, 9*(2).

Tabachnick, B. G. & Fidell, L. S. (2001). *Using Multivariate Statistics* (4th ed.). Boston: Allyn & Bacon.

Toland, M. D. & deAyala, R. J. (2005). A multilevel factor analysis of students' evaluation of teaching. *Educational and Psychological Measurement, 65*(2), 272-296.

Tomey, A. (2003). Learning with cases. *Journal of Continuing Education in Nursing, 34*(1).

Van der Hem-Stokroos, H. H., Van der Vleuten, C. P., Daelmans, H. E., Haarman, H. J. & Scherpbier, A. J. (2005). Reliability of the clinical teaching effectiveness instrument. *Medical Education, 39,* 904-910.

Von Glasersfeld, E. (1996). Introduction: aspects of constructivism. In C. T. Fosnot (ed.). *Constructivism: Theory, perspectives, and practice* (Chapter 1). New York: Teachers College Press.

Vygotsky, L. S. (1978). *Mind in society :the development of higher psychological processes.* Cambridge, MA: Harvard University Press

Waltz, C. F., Strickland, O. L. & Lenz, E. R. (2005). *Measurement in nursing and health research* (3rd ed.). New York: Springer Publishing Company.

Weimer, M. (2002). *Learner-centered teaching five key changes to practice*. San Francisco: Jossey-Bass.

Weis, P. & Guyton-Simmons, J. (1998). A computer simulation for teaching critical thinking skills. *Nurse Educator, 23*(2).

Wynd, C., Schmidt, B. & Schaefer, M. (2003). Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research, 25*(5), 508-518.

Yen, M. & Lo, L. (2002). Examining test-retest reliability. *Nursing Research, 51*(1), 59-62.

Zimmerman, L. & Westfall, J. (1988). The development and validation of a scale measuring effective clinical teaching behaviors. *Journal of Nursing Education, 27*(6).

CURRICULUM VITAE

Cynthia E. Reese

Education

Indiana University, Indianapolis, Indiana 2009

PhD in Nursing Science

- Major: Health Systems
- Minor: Nursing Education
- Focus: Effective teaching in simulations.

University of Illinois at Chicago, Chicago, Illinois 1995

Master of Science

- Emphasis: Medical Surgical Nursing

University of Cincinnati, Cincinnati, Ohio 1983

Bachelor of Science in Nursing

Professional Experience

Lincoln Land Community College, Springfield, Illinois 1999–Present

Professor of Nursing

- Responsible to teach didactic and clinical.
- Coordinator of first semester evening program.
- Teach didactic and clinical in LPN summer bridge course.

Memorial Medical Center, Springfield, Illinois 1990–2001

Clinical Nurse III – 1990-1999 Cardiac Surgery Intensive Care Unit

Staff Nurse in PRN pool – 1999-2001

Mac Murray College, Jacksonville, Illinois 1995–1999

Adjunct Professor of Nursing

Clinical Instructor in BSN program.

St. Vincent Memorial Hospital, Taylorville, Illinois 1988–1990

Staff Nurse – Intensive Care Unit

Charge Nurse – Medical Surgical Unit

Wyandotte General Hospital, Wyandotte, Michigan 1986–1987

Staff Nurse – Cardiac Care Unit

Medical Center Hospital, Chillicothe, Ohio 1983–1986

Staff Nurse – Medical Surgical Unit, Nursery, Operating Room

Awards, Certifications, and Memberships

- Fellow, Nurse Education, Illinois Board of Higher Education 2007
- Nurse Educator Certification 2005
- Nominated for Pearson Master Teacher Award 2005
- CCRN Certification 1991–2002
- Advanced Cardiac Life Support Certification 1991–1999
- Member: National League for Nursing 1999–Present
- Member: National Organization for Associate Degree Nursing 1999–Present
- Member: Sigma Theta Tau 1995–1998

Publications

- Reese, C. E., Jeffries, P. R. & Engum, S. A. (in press). Learning together: using simulations to develop nursing and medical student collaboration. *Nursing education perspectives.*

Professional Presentations

- Learning Together: Interdisciplinary collaboration between Medical and Nursing Students in Clinical Simulation 2007
  Poster Presentation, Midwest Nursing Research Society Annual Conference, Omaha, NE
- Evidence Based Teaching in Clinical Simulation (invited) 2007
  Fairbanks Simulation Institute, Indianapolis, IN

Professional Development Activities 1995–2008

- Clinical Simulations Conference, IUSON, Indianapolis, IN 2008
- Midwest Nursing Research Society Annual Conference, Indianapolis, IN 2008
- NOADN Annual Convention, Las Vegas, NV 2007
- SUN Simulation User Network Conference, Indianapolis, IN 2007
- Midwest Nursing Research Society Annual Conference, Omaha, NE 2007
- Get Real! Using Simulation to Transform Nursing Education 2006
- Critical Test Item Writing, NAODN Conference, Bloomington, IL 2005
- National League for Nursing Educators Conference, Baltimore, MD 2005
- Midwest Nursing Research Conference, Cincinnati, OH 2005
- Lab and Diagnostic Tests Update, Springfield, IL 2002
- Helping Students Prepare for NCLEX Faculty Development Workshop, Springfield, IL 2002
- Advanced Topics in Diabetes, Springfield, IL 2000
- Pharmacology Update, Springfield, IL 2000
- Bold Aims for 2000: Clinical Update for Diabetes, CHF, and Pneumonia, Memorial Medical Center, Springfield, IL 2000

- PAR System and Critical Thinking Test Item Writing, Springfield, IL — 2000
- Influencing Nursing Practice through Research, Springfield, IL — 1999
- Cardiology Article Reviewer for Critical Care Nurse Magazine — 1997–1999
- Marquette Monitoring Mentor User Support system — 1998
- BVS BI Ventricular Support System user training — 1998
- 7 th Annual CPM National Conference, Grand Rapids, MI — 1996
- National Teaching Institute and Critical Care Exposition, New Orleans, LA — 1995
- Poster Presentation of Master's research project "Accuracy of SvO2 Monitoring" Nursing Research Conference, Memorial Medical Center Springfield, IL — 1995

<u>Institutional Service Activities</u>

Lincoln Land Community College

- Nursing Evaluation Committee Member — 2007–Present
- Nursing Resources Committee Member — 2003–Present
- Nursing Curriculum Committee Member — 2003–2005
- Admissions and Retention Committee Member — 2000–2001
- Institutional Governance Committee Member — 2002–2004
- Review health forms and immunizations — 1999–Present
- Continuing Education Committee Member — 1999–2000

Memorial Medical Center

- Divisional C Q I Committee Member — 1991–1999
- Chair of Unit Based CQI Committee — 1995–1999
- Unit Based Self Governance Council Member — 1995–1999
- Preceptor for new staff nurses — 1992–1999

Community Service Activities

- Member of Taylorville Junior High School Athletic Boosters — 1999–2002
- Assist with Christian County YMCA Capital Campaign
- Assistant Girl Scout Leader — 1993–1999