The Dissertation Committee for LeAnn Chisholm Certifies that this is the approved version of the following dissertation:

Validating Targeted Behavioral Markers for Teamwork Skill and Performance Outcomes in Simulation

Committee:

M. Terese Verklan, PhD, CCNS, RNC, FAAN, Chair

Sheryl Bishop PhD

Carol Wiggs PhD, RN, CNM, AHN-BC

Leigh Ann Cates PhD, APRN, NNP-BC, RRT-NPS, CHSE

Bill Boudreaux Ed.D, RN

Rachel Kilgore PhD, RN

_____

Dean, Graduate School

Validating Targeted Behavioral Markers for Teamwork Skill and Performance Outcomes in Simulation

by

LeAnn Jeanette Chisholm MSN, RN

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas Medical Branch

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas Medical Branch

July 1, 2016

Dedication

This work is dedicated to patients, families, and healthcare providers who endure suffering due to patient safety events. These events can result in death or disability to patients, cause family members to become caregivers, widows, and orphans, and can also destroy nurses and healthcare providers. This work is the beginning of a journey. If this and future studies of teamwork help just one patient, family, or nurse and even prevent just one error, it will be worth it.

Acknowledgements

This work would not be possible without the help of many people. I would like to acknowledge the love and support of my family, particularly my husband Chris, who continually encouraged me throughout this journey. I would also like to acknowledge several faculty members who believed in the importance of this research and encouraged me along the way: Rose Harding, Cindy Ceynar, Keili Peterman, Lisa Donnelly, Mary Ford, Kellie White, Theresa Rhodes, Dr. Gina Hale, and Jenny Thedford. I would also like to acknowledge the exceptional mentorship of my chair, Dr. Terese Verklan, who provided direction and insight. In addition, the complex analysis would not have been possible without the extraordinary guidance of Dr. Sheryl Bishop. The simulation expertise of Dr. Leigh Ann Cates helped to refine the educational program and simulation experience which was provided to the subjects. I would also like to acknowledge the encouragement provided by Dr. Rachel Kilgore, Dr. Bill Boudreaux, and Dr. Carol Wiggs along with their medical, nursing and teamwork expertise. I would like to thank all of these individuals for their patience, guidance, and support along this journey. Most importantly, thank you Lord for giving me the strength to endure, and placing all of these wonderful people in my path.

Validating Targeted Behavioral Markers for Teamwork Skill and Performance Outcomes in Simulation

Publication No._____

LeAnn Jeanette Chisholm, PhD

The University of Texas Medical Branch, 2016

Supervisor: M. Terese Verklan

The National Academy of Medicine, formerly the Institute of Medicine, asserts healthcare providers must demonstrate competency in teamwork (IOM, 2001). TeamSTEPPs® is a comprehensive instructional framework developed by the Agency for Healthcare Research and Quality and Department of Defense to teach and evaluate teamwork (King et al., 2008). TeamSTEPP's® evaluation instruments include the Teamwork Attitudes Questionnaire, Teamwork Perceptions Questionnaire, and the Teamwork Performance Observation Tool (TPOT). Assessment of teamwork perceptions and attitudes can be used to indicate the effectiveness of team training (Brock et al., 2013). The TPOT (AHRQ, 2014a) was designed to objectively measure teamwork skill performance but has not been sufficiently studied.

This methodological study seeks to demonstrate the viability of the Team Performance Observation Tool enhanced with Targeted Behavioral Markers (TBMs) as a

sensitive and valid measure of teamwork performance and a superior approach to teamwork assessment without behavioral markers. A multitrait-multimethod research design will be used with an exemplar convenience sample of baccalaureate nursing students to correlate results from the TPOT and behavioral markers instruments, teamwork perceptions, teamwork attitudes, and the National League for Nursing (NLN) Simulation Checklist for convergent validity and with the Clinical Skills Self-efficacy Scale for divergent validity. Outcomes identified further needed refinement for rigorous psychometric evaluation for the TPOT and provided mixed support for validation of specific TBMs to measure desired teamwork performance outcomes.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Abbreviations

AHRQ	Agency for Healthcare Research and Quality
ANTS	Anesthetists' Non-Technical Skills
C	Communication Subscale
CATS	Communication and Teamwork Skills
CSES	Clinical Skills Self-efficacy Scale
EBAT	Event Based Approach to Training
EBT	Event Based Training
GSBS	Graduate School of Biomedical Science
ICC	Intraclass Correlation Coefficient
IOM	Institute of Medicine (now the National Academy of Medicine)
IRB	Institutional Review Board
L	Leadership Subscale
MS	Mutual Support Subscale
MTMM	Multitrait-multimethod
NAM	National Academy of Medicine (formerly Institute of Medicine)
NCSBN	National Council of State Boards of Nursing
NLN	National League for Nursing
OTAS	Observational Teamwork Assessment of Surgery
PEARLS	Promoting Excellence and Reflective Learning in Simulation
PI	Principle Investigator
QSEN	Quality and Safety Education for Nurses
SM	Situation Monitoring Subscale

SPSS	Statistical Package for Social Sciences
TBMs	Targeted Behavioral Markers or behavioral markers
TDC	Thesis and Dissertation Coordinator
TeamSTEPPs	Team Strategies and Tools to Enhance Performance and Patient Safety
TPOT	Teamwork Performance Observation Tool
TS	Team Structure Subscale
T-TAQ	TeamSTEPPs Teamwork Attitudes Questionnaire
T-TPQ	TeamSTEPPs Teamwork Perceptions Questionnaire
UTBMAF	University of Texas Behavioral Marker Audit Form
UTMB	University of Texas Medical Branch

Chapter One Introduction

INTRODUCTION

Patient safety is a top priority for nurses and healthcare providers alike, yet medical errors are not decreasing. The Agency for Healthcare Research and Quality (AHRQ, 2014f) indicates that one out of seven Medicare patients will be the victim of a medical error. The National Academy of Medicine's report, formerly the Institute of Medicine (IOM, 2001), reported that 98,000 deaths per year were attributed to medical errors. A more recent study concluded that medical error accounts for 200,000 to 440,000 deaths per year (James, 2013), with many of these medical errors related to ineffective teamwork and communication (AHRQ, 2014f; IOM, 2001). In an effort to improve teamwork and communication in healthcare, the Department of Defense and AHRQ developed the Team Strategies and Tools to Enhance Performance and Patient Safety (TeamSTEPPs®) Instructional Framework to facilitate teamwork education and evaluation. The two government agencies encourage nationwide implementation of these essential teamwork concepts in order to decrease morbidity and mortality associated with medical error.

THE BACKGROUND OF THE STUDY

The National Academy of Medicine (NAM), Quality and Safety Education for Nurses (QSEN), and The Joint Commission all recommend educating healthcare providers in team situations in order to improve patient safety and outcomes (IOM, 2001; Joint Commission, 2008; QSEN, 2012). Teamwork has been identified as an essential competency by the NAM (IOM, 2003) and QSEN (2012). Despite strong assertions from these national agencies, reports of objective teamwork assessment and evaluation in undergraduate nursing programs are limited. Smith et al. (2007) explored whether nursing programs include the QSEN teamwork

competencies while Barton et al. (2009) examined the appropriate level to introduce the QSEN competencies into nursing curricula, and MacDonald et al. (2010) assessed students' and professionals' perspectives on the essential competencies. QSEN held institutes from January 2010 to November 2011 to provide support for faculty in implementing the competencies. During these institutes, 662 faculty were surveyed about inclusion of the QSEN competencies in their respective nursing curricula. According to respondents in this survey, Teamwork and Collaboration were the least integrated competencies (Barnsteiner et al., 2012). Cronenwett et al. (2007), asserts that didactic instruction or a single course addressing the QSEN competencies is not enough. The competencies need to be reinforced in clinical experiences, simulations, and skills laboratories (labs). However, simulations and clinical labs were reportedly the least utilized methodology to teach and evaluate QSEN competencies (Sullivan et al., 2009). From these studies, it may be ascertained that the QSEN teamwork competencies are important and are included to some extent in nursing curricula; however, objective observational performance outcomes of teamwork are not routinely reported.

STUDY PROBLEM

The environment surrounding nursing practice is complex and continually evolving. Preparing future nurses to function in this environment is a challenge; therefore, nursing education needs to evolve to meet the challenge (Benner et al., 2010). Healthcare providers, particularly nursing students, need to practice teamwork in a safe learning environment and receive specific feedback in order to improve performance (Billings & Halstead, 2012). Simulation provides an avenue with which to place nursing students in realistic and replicable clinical situations that require teamwork. There are endless clinical situations that can be replicated through simulation. However, in order to evaluate teamwork performance accurately

and provide constructive feedback to students, reliable evaluation instruments are needed. In addition, these newly developed instruments need to be tested using a reliable and valid methodology.

RESEARCH QUESTION AND AIM OF STUDY

The purpose of this study is to analyze psychometric properties of the Teamwork Performance Observation Tool (TPOT) enhanced with Targeted Behavioral Markers (TBM) using a methodology by which validation can be efficiently accomplished and replicated with various teamwork simulation scenarios.

The first aim of this research is to demonstrate the viability of the Team Performance Observation Tool enhanced with Targeted Behavioral Markers as a sensitive and valid measure of teamwork skill and performance acquisition in simulation training. Three hypotheses will be tested:

A1H1: Interrater agreement on simulation performance will be higher for the behavioral markers (TBMs) than the TPOT demonstrating higher rating reliability and specificity.

A1H2: There will be a significant difference in behavioral marker assessments from TPOT assessments post simulation training demonstrating better performance assessment with the behavioral markers.

A1H3: Correlations between teamwork performance evaluations (TPOT and behavioral markers) with TeamSTEPPs® attitudes and perceptions will be higher for the behavioral markers than the TPOT demonstrating greater sensitivity of the behavioral markers.

The second aim of the study is to assess the convergent and divergent validity of the behavioral markers as a valid measure of skill and performance acquisition in simulation training. Two research questions will be addressed:

RQ 1: What is the relationship between the original TPOT, behavioral markers, teamwork perceptions, teamwork attitudes, and NLN Simulation Checklist (assesses convergent validity)?

RQ2: What is the relationship between the original TPOT, behavioral markers, and the Clinical Skills Self-efficacy Scale (assesses divergent validity)?

SIGNIFICANCE

Nursing students learn about the science of nursing in the classroom and are expected to apply it in the clinical setting. Some clinical skills, such as teamwork, may only be observed by nursing students in the clinical setting, particularly during critical events when patients' lives are at stake. It takes time and exposure to an environment in order to 'fit in' and become a contributing member of the team (Malouf and West, 2011). In order to facilitate teamwork improvement, nursing students need to have the opportunity to reflect on teamwork and receive feedback on how to improve (Hirokawa, 2012). The opportunity to improve teamwork is critical to student learning and needs to occur prior to entering the high stakes healthcare workforce. Benner et al. (2010) calls for nurse educators to radically transform nursing education by providing more experiential learning opportunities which facilitate students' development of clinical judgment.

OVERVIEW OF THE RESEARCH METHODOLOGY

This methodological study seeks to demonstrate the viability of the Team Performance Observation Tool enhanced with Targeted Behavioral Markers (TBMs) as a sensitive and valid

measure of teamwork performance and a superior approach to teamwork assessment without behavioral markers. A multitrait-multimethod research design was used with an exemplar convenience sample (n=54) of baccalaureate nursing students to correlate results from the TPOT and behavioral markers instruments, teamwork perceptions, teamwork attitudes, and simulation intervention checklist for convergent validity and with the Clinical Skills Self-efficacy Scale for divergent validity. Outcomes identified further needed refinement for rigorous psychometric evaluation for the TPOT and provided mixed support for validation of specific behavioral markers (TBMs) to measure desired teamwork performance outcomes.

DEFINITION OF RELEVANT TERMS

Team structure "facilitates teamwork by identifying the individuals among which information must be communicated, a leader must be clearly designated, and mutual support must occur" (AHRQ, 2014e, B-2-7).

Leadership involves "identifying a goal and defining a plan to achieve the goal, assigning tasks and responsibilities, sharing the plan, monitoring the plan and progress toward the goal, modifying the plan and communicating changes to all team members, and reviewing the team's performance" (AHRQ, 2014e, B-4-9).

Situation Monitoring "is a way for team members to be aware of what is going on around them. This awareness will enable individuals to adapt to changes in the situation and will also create opportunities to support other team members when needed. Situation monitoring is moderated by communication, which allows for the sharing of new and emerging information with other team members, to develop and maintain a shared mental model" (AHRQ, 2014e, B-5-8).

Mutual Support "involves team members 1) assisting one another; 2) providing and receiving feedback; and 3) exerting assertive and advocacy behaviors when patient safety is threatened" (AHRQ, 2014a, B-6-7).

Communication "is a process whereby information is clearly and accurately conveyed to another person using a method that is known and recognized by all involved. It includes the ability to ask questions, seek clarification, and acknowledge the message was received and understood. One critical result of effective communication is a shared understanding, between the sender and receiver(s) of the information conveyed" (AHRQ, 2014e, B-3-10).

Clinical judgment is "…case-based, contextually bound, interpretive reasoning. It is always in the context of the particular patient…simulation is an example of case-based teaching, providing students with real-time experience of thinking in action" (Billings & Halstead, 2012, p. 28).

THEORETICAL FRAMEWORKS

Three theoretical frameworks guided the development and implementation of this study: TeamSTEPPs®, the theory of nomological networks, and the Nursing Education Simulation Framework. TeamSTEPPs® is a teamwork instructional framework for healthcare providers. The framework includes educational tools and strategies to teach the identified competencies and instruments for measuring teamwork attitudes, perceptions, and performance. The theory of nomological networks was used to create the Teamwork Nomological Network. When nomological network theory was applied to the concept of teamwork and established teamwork measurement instruments, the Teamwork Nomological Network was created. The Nursing Education Simulation Framework was used to guide the development and implementation of the teamwork simulation experience in this study.

TEAMSTEPPS® INSTRUCTIONAL FRAMEWORK

Team Strategies and Tools to Enhance Performance and Patient Safety (TeamSTEPPs®), is a comprehensive teamwork instructional framework developed by the US government to teach and evaluate teamwork (King et al., 2008). Twenty-five years of research, including evidence from aviation, influenced the Agency for Healthcare Research and Quality and the Department of Defense in the development of TeamSTEPPs® for healthcare providers (King et al., 2008). The framework consists of evidence-based concepts which can be used to teach and evaluate teamwork. Therefore, the framework includes strategies for teaching the competencies, tools to facilitate application, and instruments to evaluate achievement of the competencies. The educational strategies and measurement instruments developed for TeamSTEPPs® are all based on the same five concepts. Outcomes of the patient care team can be measured by teamwork attitudes, knowledge, and performance (King et al., 2008, see Figure 1, TeamSTEPPs® Instructional Framework).

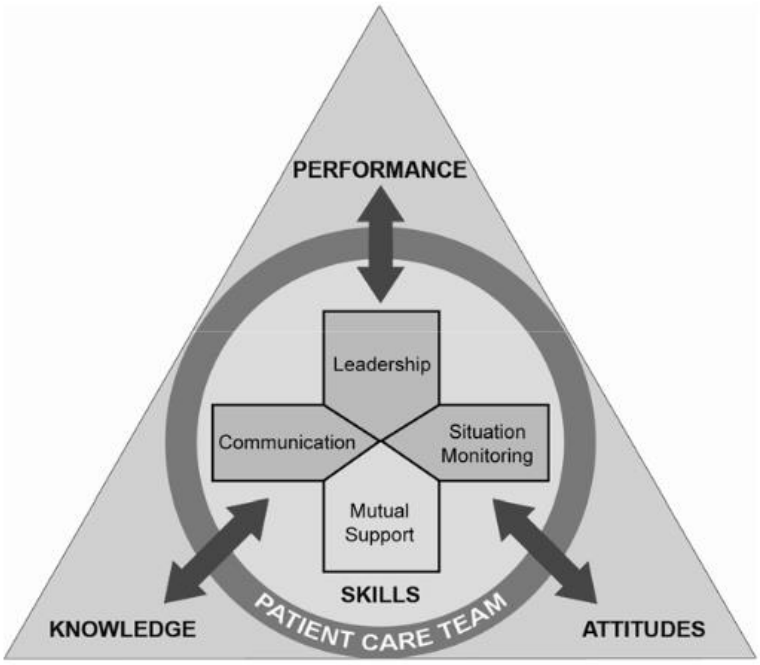


Figure 1, TeamSTEPPS® Instructional Framework

The core concepts of TeamSTEPPs® include team structure, leadership, situation monitoring, mutual support, and communication. Team structure is the identification of members who constitute the team. The team should clearly identify a leader and lines of communication (AHRQ, 2014e, B-2-7). The second concept is Leadership. A specific leader is needed to provide direction and monitor the progress of the team (AHRQ, 2014e, B-4-9). Situation Monitoring involves an awareness and shared understanding among team members. This is necessary for team members to accomplish a shared goal and improvement of the patient situation (AHRQ, 2014e, B-5-8). Mutual Support is the support individuals on the team provide to one another in order to accomplish a shared goal (AHRQ, 2014a, B-6-7). The last and possibly most important concept in teamwork is Communication. It is critical for members of the team to share information. Effective communication is characterized by the sending and receiving of information accurately (AHRQ, 2014e, B-3-10).

TeamSTEPPs® evaluation instruments include the Teamwork Attitudes Questionnaire, Teamwork Perceptions Questionnaire, and the Teamwork Performance Observation Tool (TPOT). These instruments were developed as a part of TeamSTEPPs®. All three instruments consist of the same five subscales: Team Structure, Leadership, Situation Monitoring, Mutual Support, and Communication. These instruments assist in the evaluation of learning after TeamSTEPPs® education (AHRQ, 2014d; King et al., 2008). Assessment of teamwork perceptions and attitudes can be used to indicate the effectiveness of team training (Brock et al., 2013). The TPOT (AHRQ, 2014a) was designed to objectively measure teamwork skill performance but the developers encourage investigators to individualize the instrument to the situation it is intended to evaluate.

The TeamSTEPPs® Instructional Framework identifies several barriers to teamwork both at the individual and team levels (King et al., 2008). The framework provides tools and strategies to overcome these barriers and improve teamwork. Improved teamwork knowledge, attitudes, performance, and patient outcomes are expected results of this process. In addition, the framework provides a guideline for organizational safety culture assessment, developing an action plan to promote change, and reevaluation of the organizational culture and outcomes (King et al., 2008).

Currently, a gap in the literature remains regarding the interrelationships between teamwork attitudes, perceptions and performance as measured by the TeamSTEPPs® Teamwork Attitudes Questionnaire (T-TAQ), TeamSTEPPs® Teamwork Perceptions Questionnaire (T-TPQ), and the TeamSTEPPs® Teamwork Performance Observation Tool (TPOT). This study used tools and strategies from the TeamSTEPPs® instructional framework to teach teamwork to senior level baccalaureate nursing students. Teamwork attitudes, perceptions, and observed performance were measured using TeamSTEPPs® instruments. Data from the instruments employed in this study were used to analyze the theoretical relationships between teamwork attitudes, perceptions, and performance and compare them with newly developed and specific targeted behavioral markers for teamwork.

NOMOLOGICAL NETWORK THEORY

According to Cronbach and Meehl (1955) a nomological network is an "interlocking system of laws which constitute a theory" (p. 290) involving a construct and the potential relationship to other constructs. The laws governing a nomological network may be observable measures/tests or theoretical in nature and relate to the same construct or relate one construct to another. The relationship between different types of observations or patterns of observations is

visually displayed in a nomological network illustration (Pedhazur & Schmelkin, 1991). The relationship of specific tests to a construct may be proposed and subsequently adopted if there is supporting evidence. Evidence of construct validity is established when theorized relationships are supported with data. The Teamwork Nomological Network (Figure 2) developed for this study proposes that relationships exist between measures of teamwork in this study, individual traits, and teamwork outcomes. In this study, the Teamwork Nomological Network illustrates how established instruments may theoretically relate to new instruments and contribute to the analysis of convergent and divergent validity using a Multitrait-Multimethod Correlation Matrix (Appendix B). If multiple methods or instruments are measuring the same construct, one would expect the correlations to be high, thereby demonstrating convergent validity (Portney & Watkins, 2009). Similarly, correlations with different traits are expected to be low and demonstrate divergent validity (Portney & Watkins, 2009).

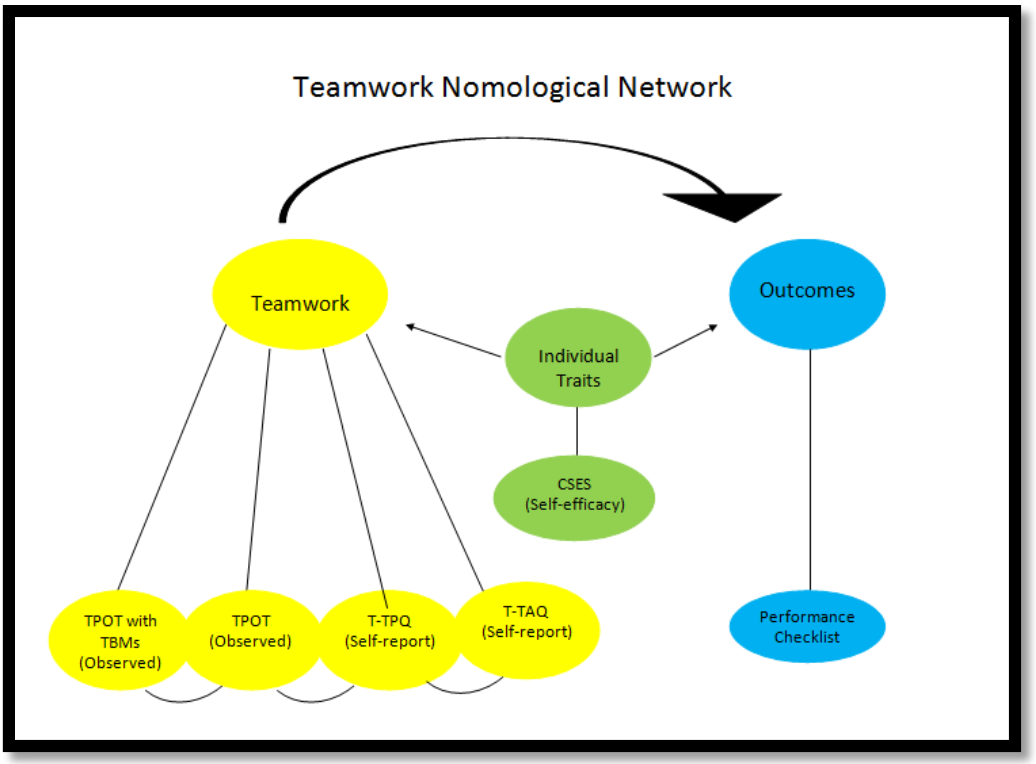


Figure 2, Teamwork Nomological Network

The behavioral markers instrument (TBM) is intended to measure observed teamwork performance. Twenty-five years of literature from aviation and the military supports a theoretical relationship between teamwork skill acquisition, teamwork performance, teamwork perceptions, teamwork attitudes, and patient/simulation outcomes (King, et al.). The relationships between the concepts in the literature provide a foundation for the Teamwork Nomological Network.

Social Learning Theory (Bandura, 1986) asserts that self-efficacy is an individual trait which leads to improved performance. Pike and O'Donnell (2010) engaged pre-registration nursing students in simulation experiences. While these students demonstrated increased self-efficacy in communication after the experience, the relationship between self-efficacy and performance was not investigated. Robertson and Felicilda-Reynaldo (2015) compared information literacy confidence levels of entry level family nurse practitioner students with their actual performance on an information literacy skills test. Although students demonstrated high levels of confidence, performance on the skills exam was poor. In addition, Liaw, Scherpbier, Rethans, and Klainin-Yobas (2012) examined the relationship between knowledge, self-efficacy, and performance. In this study, neither knowledge nor self-efficacy was a predictor of clinical performance. Porter, Gogus, and Yu (2011) assert that collective self-efficacy beliefs can mediate the disappointing effects of not achieving team performance goals; however, the specific relationship between self-efficacy and team performance is not clear. Oetker-Black and Kreye (2015) analyzed psychometric properties of the Clinical Skills Self-efficacy Scale (CSES) and assert that self-efficacy is a factor which leads to clinical skill competence. The CSES was included in the Teamwork Nomological Network as an individual attribute. The relationship between clinical skill self-efficacy and teamwork was not clearly established; however, it was

not expected to be highly correlated to teamwork or teamwork performance. Therefore, the CSES will be used to assess for divergent validity in the MTMM Correlation Matrix.

NURSING EDUCATION SIMULATION FRAMEWORK

The Nursing Education Simulation Framework is a theoretically based model for simulation development, implementation, and evaluation which is based upon evidence from nursing, medicine, and other healthcare disciplines. A primary assumption of this framework is that student learning is cognitive, experiential, and sociocultural (Jeffries, 2007). Additionally, the framework assumes that there is a critical interplay between the teacher, student, educational practices, simulation design characteristics, and outcomes (Jeffries, 2007). Essential elements of the framework include the teacher/facilitator, active learning, diverse learning styles, collaboration, and high expectations.

According to the framework, the teacher serves as the simulation facilitator and evaluator of the student experience. Five faculty in this study were assigned three to four teams of students. The faculty role involved orienting students to the experience, facilitating the simulation, and facilitating the debriefing session after the simulation. Students need to be actively engaged in the simulation experience whether they are participating as a nurse, family member, or observer. In this study, each student portrayed a team leader, primary care nurse, or family member. Jeffries (2007) asserts the roles and expectations of the student should be clear at the beginning of the simulation. Therefore, student participants in this study received information regarding their role and an orientation to the simulator and simulation experience.

Active learning is also an essential simulation educational practice (Jeffries, 2007) that occurs when students are allowed to make decisions. In this study, students participated in a simulation involving a patient with acute pancreatitis and hemodynamic instability. Students

were asked to complete a set of critical thinking questions about acute pancreatitis prior to the simulation which promoted active learning and facilitated each group beginning the simulation with similar baseline knowledge. Students made decisions and clinical judgments during the simulation. These decisions, whether correct or incorrect, promote critical thinking (Billings and Halstead, 2012). Feedback on performance is a part of active learning; however, the teacher/facilitator needs to be careful to maintain the safe learning environment during the feedback session (Jeffries, 2007). Faculty in this study reinforced the confidential nature of simulations and debriefings in order to promote a safe learning environment.

Simulation accommodates a wide variety of learning styles. Visual learners respond to a realistic clinical environment and auditory learners are accommodated by verbal handoff communication and other verbal cues. Tactile learners appreciate the hands on approach to physical assessment of a mannequin because they can palpate a pulse as well as auscultate heart and lung sounds. In addition kinetic learners appreciate using realistic equipment such as blood pressure cuffs, glucose monitors, and cardiac monitoring equipment. In this study, students were presented with a handoff report involving a patient with acute pancreatitis. The patient in this simulation was portrayed by a Laerdal 3G® mannequin with realistic pulses, auscultation sounds, and additional elements such as a patient monitor, patient identification information, and a realistic patient chart in order to increase realism and address various styles of learning.

True collaboration between the teacher and student leads to a positive learning experience and is an essential element of the framework (Jeffries, 2007). During the debriefing, faculty provide feedback to students regarding performance and students provide valuable feedback to faculty which will help the faculty improve future simulations. This study used the Promoting

Excellence and Reflective Learning in Simulation (PEARLS) debriefing script (Eppich and Cheng, 2015) to promote student and faculty collaboration during debriefing.

Positive outcomes of simulations are more likely to occur when both student and teacher approach the simulation with high expectations (Jeffries, 2007). In this study, a simulation was used to promote student use of teamwork behaviors. Instruments were used to analyze teamwork perceptions, attitudes, and the observed performance of teamwork. However, this simulation produced additional outcomes as expressed by faculty and students which will be discussed in Chapter 5.

According to Jeffries (2007) simulation design should include objectives, fidelity, problem solving, student support, and reflective thinking (debriefing). Clear objectives guide the learning and assist in determining that learning outcomes have been met. The acute pancreatitis simulation used in this study was developed by the National League for Nursing and Laerdal and supervised by Dr. Pamela Jeffries. The simulation included objectives, a realistic scenario and patient chart documents, and a plan for debriefing. Fidelity refers to the realism of the simulation. The amount of fidelity required for the simulation depends on the objectives. Simulations requiring assessment of a pulmonary problem would require a simulator/mannequin with lung sounds. Problem solving should be tailored to the student level and capabilities. The student participants in this simulation were senior level baccalaureate students in the final semester of nursing school. The simulation required students to synthesize information, communicate with one another, and recognize the need for additional assistance since the patient was not improving with current interventions. This situation was complex but attainable.

Students also need to be supported in the learning process. Jeffries (2007) asserts that cues can be provided to students during simulation in order to support the learning process;

however, it is important that the cues do not hinder the student's problem solving process. It was not necessary to provide many cues during the simulation in this study. A Simulation Center staff person ran all of the simulations in order to reduce variability and assure that cues provided were similar across simulations. The Simulation Center staff person explicitly followed the simulation script which was provided by the NLN.

Finally, debriefing or reflective thinking about the simulation is guided by the faculty facilitator immediately after the simulation and focuses on the learner objectives. This process should be constructive rather than critical and lead the students to the correct line of logic rather than a summary of correct and incorrect behaviors (Jeffries, 2007). Faculty in this study used the PEARLS (Eppich and Cheng, 2015) debriefing script in order to promote consistency between group debriefings and address all learning outcomes.

The final element of the Nursing Education Simulation Framework involves outcomes. Knowledge, satisfaction, self-confidence, skill attainment, and critical thinking are all potentially measurable outcomes of simulation. A plan for measuring outcomes of simulation is a critical element of the framework and needs to be established prior to implementing a simulation (Jeffries, 2007). Information gained from measuring outcomes helps to improve future simulations and indicates whether objectives of the simulation have been met (Jeffries, 2007). This study used several outcome measures which will be discussed in Chapters Three, Four, and Five. In addition to instruments specific to this study, faculty and students expressed anecdotal outcomes which will be discussed in Chapter Five.

THE STUDY DELIMITATIONS

The study was conducted at one mid-sized University in Southeast Texas with senior level baccalaureate nursing students entering the Preceptorship nursing course. This is a

synthesis nursing course which consists of a 96 hour clinical experience with a nurse preceptor in the hospital setting. Data for the study was collected on three dates in October 2015 prior to students beginning their hospital experience with a preceptor. This study does not involve extensive instrumentation review or analysis of the simulation specific behavioral markers. However, methods used to develop the behavioral markers in preliminary work will be discussed. The study is limited to analysis of the interrelationships between the behavioral markers and theoretically related instruments (TPOT, TPQ, TAQ, NLN Checklist) and a theoretically unrelated instrument (CSES) in order to provide information regarding psychometric properties of the behavioral markers in this study and provide a methodology for assessment of future behavioral markers.

ASSUMPTIONS

This study has three main assumptions. First, the study assumes that teamwork behaviors can be observed and measured during simulated experiences. A second major assumption is that faculty assigned to rate team performance did so to the best of their knowledge and ability. Lastly, the study assumes that students answered individual survey questions honestly according to their attitudes and perceptions of teamwork after the simulation experience.

CONCLUSION AND ORGANIZATION OF THE CHAPTERS

Simulations developed and implemented using the Nursing Simulation Education Framework provide an avenue for educating nursing students in a safe environment, allowing them to make clinical judgments and practice communicating critical information (Jeffries, 2007) in situations which require teamwork. The debriefing process, which occurs after the simulation, encourages reflective thinking and processing information in order to improve future decision making and achieve desired outcomes (Jeffries, 2007). When simulations incorporate the

TeamSTEPPs® Instructional Framework, teamwork skills can be learned and evaluated; however, additional psychometric data regarding the instruments is needed. The behavioral markers are scenario specific teamwork behaviors; therefore, they can facilitate the evaluation of teamwork and provide focus to the post-simulation debriefing session aimed at improving teamwork. The employment of rigorous methods to assess construct validity of the behavioral markers is needed. Currently, there is limited information regarding psychometrics of the TPOT (Zhang, 2013, 2015; Maguire, Bremmer, & Yanosky, 2014) and instruments which measure simulation specific teamwork behavioral markers.

The sections of this study to follow will be organized into five chapters. In Chapter Two, the literature regarding the importance of teamwork education and known methods to evaluate teamwork will be outlined and analyzed. Chapter Three involves discussion of the research methodology, which will include the research design, sample population, recruitment, data collection and instruments, plan for data analysis, and limitations. The study results and a discussion of the findings will be presented in Chapter Four. A summary of the study findings, conclusions, and recommendations will be offered in Chapter Five followed by a complete list of references and appendices.

Chapter Two Review of Literature

Chapter Two provides a synthesis of literature regarding the historical development of teamwork education in healthcare, instruments and methods for measuring teamwork in healthcare, and evidence indicating that teamwork education translates to improved patient outcomes. Literature involving teamwork education, measuring teamwork, teamwork behavioral markers, simulation in nursing education, and translational research will be discussed. Gaps in the literature are also identified and will be discussed as a rationale for this study.

TEAMWORK EDUCATION

The aviation industry recognized long ago that errors causing airplane accidents occurred more frequently due to human factors rather than mechanical failure (Thomas, et al., 2003). Crew Resource Management is a team training strategy developed by the aviation industry to improve teamwork and promote safety (Thomas et al., 2003). Pilots are required to frequently demonstrate continued competence in these skills. As a result of team training and other initiatives, the International Civil Aviation Organization reports aviation mortality rates in 2013 were less than one third of those from 2010 (ICAO, 2014). Twenty-five years of research, including evidence from the aviation industry and the military, influenced the Agency for Healthcare Research and Quality and the Department of Defense to develop the TeamSTEPPs® Instructional Framework, a team training and evaluation strategy for healthcare providers (King et al., 2008).

Prior to the development of TeamSTEPPs®, Bowers et al. (1994) studied the importance of teamwork in military aviation in order to identify teamwork concepts which could be used in teamwork education programs. At this time in history, global measures of team performance were common but they were founded heavily on the accomplishment of tasks at the individual

level rather than the team level. Bowers et al. (1994) asserted that objective teamwork behaviors which could be quantified, such as communication frequency, may better reflect actual teamwork performance. Later, Baker et al. (1996) used teamwork behaviors identified by Prince et al. (1993) and further tailored them for team performance in the training community, fixed-wing attack (aircraft) community, and the cargo helicopter community. Participants were asked the importance of each behavior to the overall concept of teamwork. One example of the identified behaviors included 'providing assistance to other crew members' which is very similar to a concept in today's TeamSTEPPs® instructional framework known as Mutual Support. Therefore, the authors concluded that specific and targeted teamwork behaviors should be the focus of future research and that teamwork training needs to change as subjects evolve from novice to experienced personnel in the field.

The military and aviation industry share a significant commonality with healthcare. All three are entities in which error can result in significant loss of life. These organizations all strive to be highly reliable (Baker et al., 2006). Highly Reliable Organizations are characterized by low rates of error despite the high potential for and severity of the errors that may occur. Effective teamwork is an essential element of Highly Reliable Organizations; therefore, Baker et al. (2006) called upon the healthcare research community to apply a framework of teamwork and team training to the education of healthcare providers and during the same year TeamSTEPPs® was launched.

MEASURING TEAMWORK

Rosen et al. (2008) identified best practices in team performance measurement. According to Rosen et al. it is important to ground teamwork measures in theory, design measures to meet specific student learning outcomes, accurately identify competencies

performed and associate teamwork performance measures with simulated scenario events such as teamwork behavioral markers. Measurement instruments need to include behaviors which may be observed, capture the process of teamwork, identify the cause of ineffective performance and facilitate providing meaningful feedback to participants. Rosen et al. also asserts that raters must be trained on the use of the measurement instrument in order to promote reliability and that evaluation should include multiple sources of data.

Several instruments for measuring teamwork are available; however, these instruments are typically associated with evaluating a particular discipline or type of emergent situation. Rosen et al. (2010) identified four of the major instruments used to measure teamwork in healthcare: University of Texas Behavioral Marker Audit Form (UTBMAF), Anesthetists' Non-Technical Skills (ANTS), Observational Teamwork Assessment of Surgery (OTAS), and the Communication and Teamwork Skills (CATS) instrument. Not only are these scales tailored to a specific area of practice (surgery, anesthesia, neonatal resuscitation), but they also use global ratings of teamwork concepts and/or track the number of occurrences. According to Rosen et al. (2010) global ratings are sufficient for pure assessment, but do not facilitate providing feedback for performance improvement during simulated training sessions. The Event Based Approach to Training (EBAT) is scenario specific (Fowlkes et al., 1998) which means that scenario events trigger specific teamwork behavioral markers. Since the specific teamwork behavioral markers are linked to events in the scenario, it is relatively simple to determine whether the teamwork behaviors were accomplished and provide feedback to the team for performance improvement during the feedback session. Although these types of instruments may be easier to use and promote quality feedback, they are labor intensive because they must be individualized to each training scenario (Rosen et al., 2010).

TEAMWORK BEHAVIORAL MARKERS

The utilization of observable behavioral markers specific to teamwork was identified in several studies. Thomas et al. (2010) conducted a study involving team training, simulation, and the attainment of neonatal resuscitation competencies. The randomized controlled trial compared the attainment of teamwork behavioral markers with the indicators of quality: resuscitation time and attainment of items on a neonatal resuscitation checklist. Although a significant relationship was not found between team behaviors and attainment of standards of care items on a checklist, significance was observed between the presence of team behaviors and resuscitation time. Groups receiving team training with high fidelity simulation tended to have quicker resuscitation times than groups who received training through different strategies.

Two additional studies used simulation to provide team training (Patterson et al., 2013 & Riley et al., 2011) and observed behavior change after the specific interventions. Patterson et al. provided team training to 289 participants initially and 151 returned 10 months later for re-evaluation using the Safety Attitudes Questionnaire. Friedman's test demonstrated significant attitude changes for teamwork, climate, and overall teamwork scores ($p < 0.001$) among the baseline, post intervention, and 10 month re-evaluation assessments. Although several studies mention the importance of communication as an attribute of teamwork, only one study correlated communication behaviors with effective team performance. Hirokawa et al. (2012) observed that groups with high team performance also demonstrated a significant and positive relationship to communication focused on interventions. However, it is not clear how the identified teamwork behaviors correlate with other measures of teamwork.

Only two studies were located in the literature that analyzed psychometric properties of the Team Performance Observation Tool (TPOT). Zhang (2013, 2015) applied the Event Based

Training (EBT) methodology to teamwork education and developed Targeted Behavioral Markers (TBMs) which are scenario specific and reflect teamwork behaviors on the Team Performance Observation Tool (TPOT). With EBT, simulation scenarios are embedded with events which should trigger an expected TBM response. Using this methodology, the evaluator can easily determine whether the objectives of the simulation have been met and provide feedback directed toward improving future performance (Dwyer et al., 1999). In Zhang's study, the TPOT enhanced with TBMs demonstrated good interrater reliability (kappa - 0.730). Additionally, test-retest reliability of the original TPOT (i.e. without the TBMs) was acceptable (kappa- 0.70) when completed two weeks apart by the same rater. Raters using the TPOT with TBMs demonstrated weighted kappa scores of 0.73. However, comparison of scores between raters using the TBM as compared to the rater using the original TPOT was not acceptable (kappa – 0.303, 0.212). The authors concluded that the original TPOT is subjective leading to inconsistency in ratings. When using the TBMs, raters demonstrated much higher interrater reliability.

Maguire, Bremmer, and Yanosky (2014) also analyzed psychometric properties of the original TPOT. Thirty-one educators with differing levels of education and experience rated student performance on the same simulation. Cronbach's alpha was 0.965 for the instrument. The authors assert convergent validity was established by the internal consistency of evaluation scores from nurse faculty with varying levels of education (Masters' degree vs. doctorate); however, this definition of convergent validity conflicts with Trochim (2006) who asserts convergent validity is established when results correlate with findings from other reliable and valid instruments measuring the same concept. Therefore, this study aims to assess both convergent and divergent validity in order to support construct validity as Trochim (2006)

asserts. In the current study, a multitrait-multimethod (MTMM) design compares four instruments measuring teamwork with each other and an instrument measuring self-efficacy and clinical skills. The four instruments measuring teamwork are expected to demonstrate moderate to high correlations with one another which is consistent with Trochim's definition of convergent validity. These four instruments will also be compared with the Clinical Skills Self-efficacy Scale which is expected to demonstrate low correlations and illustrate divergent validity.

The behavioral markers are scenario specific teamwork behaviors; therefore, they facilitate the evaluation of teamwork and provide focus to the post-simulation debriefing session aimed at improving teamwork. The employment of rigorous methods to assess construct validity of teamwork evaluation instruments is needed. Currently, there is limited information regarding psychometrics of the TPOT (Zhang, 2013, 2015; Maguire, Bremmer, & Yanosky, 2014) and instruments which measure simulation specific teamwork behavioral markers.

Although the Maguire et al. (2014) and Zhang (2013, 2015) studies provide much needed information regarding psychometrics of the TPOT, neither provide sufficient information regarding convergent and divergent validity which is essential to support construct validity (Portney & Watkins, 2009). The current study analyzed correlations between three established teamwork instruments, the behavioral markers in this study, and self-efficacy with clinical skills. The current study provided convergent and divergent validity data needed to analyze psychometric properties of the original TPOT and behavioral markers.

SIMULATION IN NURSING EDUCATION

The NLN's Vision for Simulation in Nursing Education emphasizes the imperative for nurse educators to purposefully integrate simulation throughout nursing curricula as a mechanism to promote good clinical judgment (NLN Vision Statement, 2015). A landmark

study conducted by the National Council of State Boards of Nursing (NCSBN) supports the use of simulation in nursing programs. In the NCSBN study, educational outcomes of six hundred sixty-six students who participated in 10%, 25%, and 50% simulated experiences in lieu of traditional clinical were evaluated and compared with educational outcomes of students in traditional clinical experiences. Findings from the NCSBN study indicate that learning in simulation can be just as effective as the learning that occurs in the clinical setting (Hayden et al., 2014). Similarly, Curl et al. (2016) conducted a study with associate degree nursing students. Students participating in 50% simulated clinical experiences which were integrated throughout the curriculum demonstrated significantly higher scores on standardized end of program exams than students participating in traditional clinical experiences alone.

Teamwork principles can be incorporated in every high-fidelity simulation (Clapper & Kong, 2012). Realistic simulations can be effective in initial training and in sustaining teamwork competencies (King et al., 2008). Student learning outcomes from these simulation experiences need to be evaluated using reliable and valid tools (NLN Vision Statement, 2015) especially when these experiences replace clinical time. Numerous instruments are available, and although researchers should consistently report psychometrics of the instruments used, reliability and validity data is not always reported (Adamson et al., 2013). Observation tools can measure the application of learning in simulation; however, the ultimate goal of nursing education, regardless of the teaching pedagogy, is for students to translate their learning into their actual clinical practice (Adamson et al., 2013).

TRANSLATIONAL RESEARCH

Highly Reliable Organizations have low rates of error although the potential for error is high (Baker et al., 2006). The literature regarding how teamwork performance in simulated settings translates to actual patient care and subsequent patient outcomes is increasing. Capella et al. (2010) provided TeamSTEPPs® Education with simulation practice to trauma teams. Actual trauma resuscitations were observed prior to and after the training. The teams demonstrated significant improvement in teamwork as measured on the Teamwork Performance Observation Tool (TPOT) as well as significantly decreased times from patient arrival to Computerized Tomography scan, arrival to intubation, and arrival to the operating room. Patterson et al. (2013) provided simulation based team training to 289 participants initially and 151 participants returned 10 months later for re-evaluation using the Safety Attitudes Questionnaire (SAQ) and a modified teamwork behavioral marker instrument for neonatal resuscitation. In this facility, patient safety events/hospital errors decreased from two to three per year to none in 1000 days. These studies support the use of teamwork education and the assessment of teamwork behaviors using the TPOT and indicate that improvements in observational teamwork behaviors in simulation are associated with improved teamwork behaviors and less error in the clinical setting.

Riley et al. conducted a study to compare birth trauma rates in three community hospitals. One hospital implemented TeamSTEPPs® didactic team training while the other two facilities used TeamSTEPPs® didactic team training along with a series of simulations conducted in the hospital setting. Hospitals using the didactic and simulation training combination demonstrated a 37% improvement ($p < 0.05$) in perinatal morbidity. Outcomes of the facility using only didactic training demonstrated no significant improvement in perinatal morbidity. Riley et al.

provide evidence that educating clinicians about teamwork and allowing them to practice and improve these behaviors in simulation is associated with improved clinical team performance and patient outcomes.

SUMMARY OF CHAPTER TWO

Chapter Two discussed the historical evolution of teamwork education and evaluation from military and aviation sources to current applications in healthcare. In addition, Chapter two provided evidence that teamwork education with simulation practice is associated with improved patient outcomes and that simulation can be used effectively in nursing education. Event Based Training was also discussed as a methodology of linking specific expected teamwork behavioral markers to simulation scenario events which can facilitate the discussion of performance improvement during the debriefing session. The Teamwork Nomological Network was developed based on the theory of nomological networks. When this nomological network theory was applied to the construct of teamwork and TeamSTEPPs® teamwork measurement instruments, the Teamwork Nomological Network was created. The Teamwork Nomological Network was used in this study to address gaps in the literature which include psychometric analysis of current teamwork evaluation instruments.

PLAN FOR REMAINING CHAPTERS

Chapter Three will discuss the research methodology for this study. Analysis and the study findings will be presented in Chapter Four. Discussion, implications of the research findings in this study, areas of future research, and conclusions will be presented in Chapter Five.

Chapter 3 Research Design

Chapter Three will outline and describe the research design. The study sampling procedures, setting, recruitment, study protocols, instruments, and rater training will be described in detail. The study aims and hypotheses will be reviewed along with procedures that were used to confidentially manage and analyze the data. Chapter Three concludes with a summary and plan for the remaining chapters.

AIMS AND RESEARCH QUESTIONS

The study had two primary aims. The first aim of the research was to demonstrate the viability of the Team Performance Observation Tool (TPOT) enhanced with Targeted Behavioral Markers (TBM) as a sensitive and valid measure of teamwork performance assessment in simulation training. The second aim was to assess the convergent and divergent validity of the Team Performance Observation Tool (TPOT) enhanced with Targeted Behavioral Markers (TBMs) as a valid measure of teamwork skill and performance acquisition in simulation training.

AIM 1, HYPOTHESIS 1

Interrater agreement on simulation performance will be higher for the behavioral marker instrument (TBMs) than the original TPOT demonstrating higher rating reliability and specificity.

AIM 1, HYPOTHESIS 2

There will be a significant difference in behavioral marker from the original TPOT demonstrating better performance assessment.

AIM 1, HYPOTHESIS 3

Correlations between performance evaluations (TPOT with and without TBMs) with TeamSTEPP® attitudes and perceptions will be higher for the behavioral marker instrument (TBMs) than the original TPOT demonstrating greater sensitivity.

AIM 2, RESEARCH QUESTION 1

What is the relationship between original TPOT, behavioral markers, and teamwork perceptions, attitudes, and the NLN Simulation Checklist (assesses convergent validity)?

AIM 2, RESEARCH QUESTION 2

What is the relationship between original TPOT, behavioral markers, and the Clinical Skills Self-efficacy Scale (assesses divergent validity)?

METHODS

The research design will be described along with methods that were used to select the study sample, recruit study subjects, and the setting in which the study was conducted. Specific study procedures such as the teamwork education program, instruments, and rater training are explicitly detailed.

RESEARCH DESIGN

This methodological study used a quasi-experimental Multitrait-Multimethod (MTMM) design to investigate construct validity of the behavioral markers instrument. The MTMM methodology was selected for its' rigorous mechanism for analyzing construct validity (Trochim, 2006). It compares methods of evaluation as well as the traits being evaluated to enable the researcher to determine if the evaluation method affected the outcome of testing. In the study, instruments with established reliability (TPOT, T-TPQ, T-TAQ) and varying methods of

evaluation were utilized to evaluate teamwork along with newly developed behavioral markers instrument which facilitated the identification of scenario specific teamwork behaviors.

The specific aims and related research questions and hypotheses were attainable using MTMM methodology. Evidence of convergent validity would be demonstrated if scores on the original TPOT, behavioral markers, positively correlate (moderate to high) with scores on teamwork perceptions and attitudes surveys, and simulation interventions as measured by the NLN developed checklist specific to the simulation. A positive correlation would mean that as TPOT and behavioral markers scores rise, perceptions, attitudes, and checklist scores should also increase. Self-efficacy has not been found to be predictive of performance (Liaw et al., 2012; Robertson & Felicilda-Reynaldo, 2015). Therefore, relationships between the original TPOT, behavioral markers, perceptions, and attitudes should have a low correlation with clinical skill self-efficacy scores. This pattern of relationships would provide evidence for divergent/discriminant validity.

STUDY SAMPLING AND SUBJECT INCLUSION CRITERIA

The study used a convenience sample of senior level baccalaureate nursing students enrolled in the Preceptorship course (n=57) at a medium sized University in Southeast Texas. Preceptorship is the final clinical capstone course, and students must successfully complete all coursework in order to enroll in the Preceptorship course. Therefore, students were eligible to participate in the research study if they successfully completed all course work and were enrolled in the Preceptorship course. No student enrolled in the Preceptorship course was excluded from the study since all students were expected to participate in the preceptorship clinical experience and successfully complete the nursing program. Participation in the research study was voluntary; however, participating in the teamwork simulation experience was a course

requirement. Fifty-six students agreed to participate; however, one student was on maternity leave which left one team of only two people. Data for this partial group was not collected for the study although the group did participate in the teamwork education, simulation, and debriefing. The total sample for the study included 54 students (n=18 teams).

SETTING

The setting for data collection was a 10 bed high-fidelity simulation center at the University. The University has 15,000 on campus students with approximately 300 on campus undergraduate baccalaureate nursing students. The Simulation Center within the Department of Nursing has ten individual patient rooms which closely resemble hospital rooms. Each room is equipped with a patient bed, high-fidelity mannequin, functioning suction and oxygen units, patient monitor, call light system, emergency call system, telephone, charting system, sink, two fixed cameras, one pan tilt zoom camera, and microphone. Faculty facilitate simulations from the control room which is located down the hall from the patient rooms. Faculty have the ability to view student interactions with the mannequin via a video server which integrates audio, video, and mannequin monitor information. Faculty control the mannequin, serve as the mannequin's voice and the physician or nurse practitioner's voice, and observe the students' performance in real time on video monitors in the control room.

In addition to patient rooms, four separate rooms are available for comprehensive video debriefing. The debriefing rooms are equipped with conference tables and a large screen monitor for viewing recorded simulations and providing performance feedback. The debriefing rooms were used for pre-briefing before the simulation, completing surveys, and debriefing after the simulation.

RECRUITMENT

The PI visited the Preceptorship course after Institutional Review Board (IRB) approval was obtained from UTMB (Appendix A) as the primary research institution. The PI explained the purpose and design of the study, answered questions, invited students to participate, and obtained their voluntary written consent (Appendix B). Fifty-six out of fifty-seven students in the cohort consented to participate and signed a written consent form. One student was on maternity leave, therefore, consent was not obtained from this student and the student's data was not collected for the study.

PROCEDURES

Fall 2015 simulation dates for scenario implementation and data collection were identified during the Summer of 2015. Faculty who were willing to conduct the teamwork simulations and were available on the identified dates, including the faculty workshop, were recruited to collect observational data, facilitate student completion of survey data, observe and debrief the team simulations. Students were randomized into nineteen teams of three students. One team consisted of only two people, therefore, data was not collected for this group or the student on maternity leave.

Due to limited faculty and student time, the one hour TeamSTEPPs® didactic teamwork education was delivered using lecture format with the entire cohort of students. Each team of nursing students participated in a simulation scenario five days to two weeks after receiving the team training. This schedule was necessary due to limited student and faculty time.

The one-hour team training for students involved a brief overview of the five key principles in TeamSTEPPs®: team structure, communication, leadership, situation monitoring, and mutual support. Specific examples of strategies which support each principle were discussed.

Short video clips from the TeamSTEPPs® website were used as exemplars to illustrate the strategies in action. Appendix C illustrates the teaching outline for the one-hour team education for nursing students in the study.

Students were informed prior to the simulation day that the simulation scenario would involve a patient with pancreatitis. Students were given a set of pre-simulation critical thinking questions focused on assessments, interventions, and evaluation strategies appropriate for use with the patient having pancreatitis. By encouraging preparation, performance issues related to knowledge deficits should have been minimized.

The simulation scenario was developed by the National League for Nursing (NLN) in collaboration with Laerdal®, a medical simulator company. The department's Simulation Specialist was trained to run the simulator, portray the physician, and the patient's voice during the simulation. The Simulation Specialist was also trained to strictly adhere to the simulation script as written in order to ensure standardization and decrease any variability in scenario implementations. Utilization of the Simulation Specialist in this capacity allowed the faculty raters to focus on evaluation, score the three instruments, and prepare for immediate debriefing of the group. Faculty observed teamwork performance and scored each team using the original TPOT, behavioral markers instrument, and the NLN Simulation Checklist. Prior to the simulation debriefing, participating students completed the demographic questionnaire, Clinical Skills Self-efficacy Scale, the teamwork perceptions (T-TPQ) and teamwork attitudes (T-TAQ) based on their team's performance during the simulation. During the debriefing, faculty led students in reflecting on how teamwork concepts and strategies were used. Videos of each team's performance during the simulation were available to illustrate areas where students successfully used teamwork concepts and areas for improvement. The total time commitment for

each student group, including teamwork education, was two and a half hours. The total time commitment for each faculty member was approximately ten to twelve hours.

Three TeamSTEPPs® Master Trainers viewed the videos of student performance and rated teams using the original TPOT, behavioral marker instrument, and the NLN Simulation Checklist. Two of the TeamSTEPPs® Master Trainers rated nine teams for a total time commitment of approximately ten hours. The PI, also a TeamSTEPPs® Master Trainer, viewed all eighteen videos for an approximate time commitment of 20 hours. Ratings from the five faculty members, PI, and two TeamSTEPPs® Master Trainers were compared and interrater agreement was examined for the study sample.

RATER TRAINING

In order to prepare faculty for the simulation, conduct interrater agreement, and develop a training video for future raters, three baccalaureate nursing students who were not in the study sample were recruited to participate in testing the simulation scenario prior to the study. These volunteers received the one-hour teamwork education training which was identical to the education received by the study subjects. The volunteers were also informed that the simulation would involve a patient with pancreatitis and were encouraged to review this concept and related patient care prior to the simulation scenario. These volunteer students did not complete any surveys or demographic forms since the experience was specifically for scenario testing and development of a training video. The volunteer student performance was realistic and revealed technical issues in the scenario implementation, pre-scenario instructions, and checklists which resulted in procedural changes made prior to study implementation. For example, the volunteers stated that a specific list of critical thinking questions related to the content needed to be provided in order to help prepare subjects for the simulation. In addition, the volunteer students

also stated that the pre-simulation instructions needed clarification. Based on these suggestions, changes were made accordingly. The volunteer students signed a confidentiality and photo release agreement prior to the simulation experience.

The researcher used the recording of the volunteer simulation scenario to refine the observational instruments and train faculty on observing teamwork behaviors. The video recording of the simulation was used during the faculty training workshop to enhance interrater agreement among the five faculty members, two TeamSTEPPs® Master Trainers, and the PI. During the workshop, faculty received an overview of TeamSTEPPs® and the instruments to be used in the study. All faculty viewed the video of the volunteers completing the simulation scenario and independently rated team performance using the original TPOT, behavioral marker instrument, and the NLN Simulation Checklist. Afterward, the researcher led the faculty raters in discussing each item in order to promote a shared understanding of the rating system and enhance interrater agreement. Interrater agreement of 80% was established prior to the study. During the workshop, the faculty raters and the Simulation Specialist rehearsed implementation of the simulation.

INSTRUMENTS

The TeamSTEPPs® Teamwork Perceptions Questionnaire (AHRQ, 2014b) was used to measure teamwork perceptions. The TeamSTEPPs® Teamwork Attitudes Questionnaire (AHRQ, 2014c) measured attitudes toward teamwork, and outcomes of the simulation were measured by the accomplishment of desired interventions on the NLN Simulation Checklist (convergence). The Clinical Skills Self-Efficacy Scale was used to assess self-efficacy with clinical skills as an individual trait since there is a theoretical relationship between individual performance and teamwork, but there is no evidence that self-efficacy with clinical skills is predictive of team

performance (divergent construct). Team overall scores were constructed for the TeamSTEPPs®, Teamwork Perceptions Questionnaire, TeamSTEPPs® Teamwork Attitudes Questionnaire, and the Clinical Skills Self-Efficacy Scale by combining individual scores for each member of the team. The same teams were evaluated using the teamwork observational instruments for comparative analyses.

TEAMWORK PERFORMANCE OBSERVATION TOOL. The original TPOT (AHRQ, 2014a) is an observational instrument that was used to evaluate teamwork behaviors during actual or simulated events (Appendix D). The instrument consists of five subscales: team structure, communication, leadership, situation monitoring, and mutual support. Each subscale consists of four to six behaviors which are rated on an ordinal scale from 1 (very poor) to 5 (excellent). The sum of each behavior in the subscale leads to a subscale score. All five subscales were summed for an overall raw score.

BEHAVIORAL MARKERS (TBMS). In preliminary work, the behavioral markers (Appendix E) were developed specifically for a NLN simulation involving a patient with Pancreatitis and hemodynamic instability. Using Zhang's (2013, 2015) methodology, specific triggering events in the simulation were associated with anticipated teamwork behavioral markers (TBMs). The situation specific behavioral markers were then linked with essential teamwork behaviors on the original TPOT. Therefore, faculty were able to follow the unfolding scenario and evaluate whether the expected behaviors were met. The behavioral markers are listed in table format along with the chronological simulation events and associated TPOT behaviors for ease of evaluation. Scores for the behavioral marker instrument indicate the overall number of teamwork behaviors which were observed.

Prior to implementation of the behavioral marker instrument, the behavioral markers were reviewed by two experts in medical-surgical nursing and two experts in TeamSTEPPs® team training methodology. The items were revised as needed to obtain basic content validity (80% agreement). Although content validity is the weakest form of validity (Polit & Beck, 2010), it is an essential first step prior to using the instruments and performing more rigorous reliability and validity testing.

TeamSTEPPS® Teamwork Perceptions Questionnaire (T-TPQ). The teamwork perceptions survey (Appendix F) is a 35 item instrument consisting of five subscales, each having seven items, which measured the perceptions of teamwork on a five point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree). The subscales measure the constructs of team structure, leadership, situation monitoring, mutual support, and communication, all of which are discussed in the TeamSTEPPS® teamwork education curriculum. The Agency for Healthcare Research Quality (AHRQ, 2010) reports Cronbach's alpha for each of the subscales ranged from 0.88 to 0.95 and construct inter-correlations were from 0.57 to 0.79.

TeamSTEPPS® Teamwork Attitudes Questionnaire (T-TAQ). The T-TAQ (Appendix G) is a 30 item instrument consisting of five subscales with six items in each subscale. The instrument measured attitudes toward teamwork on a five point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree), using the same five subscales as the perceptions survey (team structure, leadership, situation monitoring, mutual support, and communication). Each of the six items within the subscales closely resembles items on the perceptions survey, only the questions are phrased to measure attitudes about the importance of each item rather than perceptions. Cronbach's alpha for the subscales ranged from 0.70 to 0.83 with inter-item correlations ranging from .541 - .633.

NLN SIMULATION CHECKLIST. The simulation checklist was a list of assessments, interventions, and evaluations that were expected during the simulation (Appendix H). The checklist was developed by the NLN specifically for the simulation in order to evaluate student attainment of the simulation objectives. Items are dichotomously scored as observed or not observed. Scores for the NLN Checklist consisted of a total number of items completed on the instrument. No reliability data is available for this instrument. Content validity was established with two medical-surgical nurse experts prior to using the instrument.

CLINICAL SKILLS SELF-EFFICACY SCALE (CSES). The CSES (Oetker-Black & Kreye, 2015) is a nine item instrument which measures student confidence with clinical skills on a zero to ten scale (0 = no confidence to 10 = total confidence) (Appendix I). The CSES consists of two subscales. This instrument measures self-efficacy ratings of hands on clinical skills such as intramuscular injection, sterile dressing change, Foley catheter insertion, nasogastric tube insertion, initiating intravenous access, calculating medication dosages, and transferring an immobile patient to a chair. The instrument has a Cronbach's alpha of 0.84. The sum of all nine items was used to create an overall raw score for each individual and team. The CSES was used to demonstrate divergent validity. The CSES was not expected to correlate highly with measures of team performance or outcomes.

DATA MANAGEMENT

The plan for data management included measures to maintain security, confidentiality, and accuracy of data. Data consisted of individual student surveys (Demographics, Clinical Skills Self-efficacy Scale, perceptions, attitudes) as well as group team performance evaluations (original TPOT, behavioral markers, NLN Simulation Checklist). Each student was assigned an individual identification number and a team identification number. The faculty evaluator for

each team was also coded. A code book linking subjects with their individual, team, and faculty evaluator identification numbers was maintained on the PI's password protected laptop computer and in a locked office. Videos of each simulation were saved on an external drive and maintained in a locked desk within the PI's locked office. Surveys and team performance evaluations included individual and team identification numbers in order to maintain confidentiality of data. Student attendance for the simulation was maintained on a course roster to verify they completed the teamwork simulation course requirement. Student completion of the required activity was reported to course faculty.

Data was collected from subjects using the survey format on Blackboard® which was associated with the preceptorship course. Since the surveys were completely anonymous, the subjects were asked to provide their individual and team code number at the beginning of the survey. Blackboard® surveys are also equipped with an option to prevent skipping items which were employed for the student subjects who were participating in the research. However, using the 'force completion' option in Blackboard® did not prevent the students from skipping items. Demographic data was coded using numbers indicating a category (example: male/female, ethnicity, etc.). Survey data from the perceptions and attitudes were recorded from 1 to 5 (1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree). Scores on the TPOT were documented as a raw score/sum of each item for each of the subscales and an overall summed raw score. Scores on the behavioral marker instrument and the NLN Checklist were documented as the overall raw score of behaviors observed. Clinical Skills Self-Efficacy Scores were identified by each subject and were summed to create an overall raw score. Faculty completed paper and pencil TPOT, behavioral marker, and NLN Simulation Checklist during the scenario. The PI entered data from the faculty forms into an Excel spreadsheet that was

imported into the Statistical Package for Social Sciences (SPSS) statistical package for the analyses. In addition to individual scores on the perceptions, attitudes, and skill self-efficacy (CSES), team mean scores on all three instruments were compiled to mirror the same set of subjects as the simulation teams in order to compare with the team observational scores (TPOT, behavioral markers, NLN Checklist). Faculty were instructed to encourage subjects to complete all surveys. Missing data was observed, despite using settings in Blackboard® to prevent it. Thirty-three individual items were missing on the demographic variables of education, healthcare experience, and experience with team training. Sixteen individual items were missing on the T-TPQ (perceptions), eight items were missing on the T-TAQ (attitudes), and one item was omitted on the CSES. Missing data was left blank and blank items were coded as missing in SPSS. The code for missing data was documented in the code book. The code book was used to link variables and labels with brief descriptions or meanings.

DATA ANALYSIS

Subject survey data obtained through the Blackboard® survey format and paper and pencil data from faculty assessment of teams was analyzed using the SPSS version 22. Two data sets were initially created since data were collected at the individual level through subject surveys and at the team level, provided by faculty rating each team. A third data set was created for team level data on surveys and observational instruments. Subject survey data was summed to mirror the team assignments and provide a team level of survey data suitable to compare with team level observation data provided by faculty. In addition, rater scores were summed for each team in order to create team scores on the observational instruments. Preliminary analysis was conducted in order to identify differences in groups on the demographic variables. After preliminary analysis, each hypothesis and research question was analyzed. Interrater Agreement

was calculated using Intraclass Correlation (ICC), and then data were analyzed according to each hypotheses and research question.

PRELIMINARY ANALYSIS

Descriptive analysis of all demographics and variables was conducted first. Data was compiled for a total of 18 teams which consisted of 54 subjects. Individual level subject data were collected from the demographic form, attitudes, perceptions, and skill self-efficacy (CSES) surveys. Team level data were collected from a total of eight faculty using the TPOT, behavioral markers, and NLN Checklist. Raw/summed scores on the subscales were calculated for the attitudes and perceptions instruments using SPSS. Overall raw/summary scores were calculated in SPSS for the skill self-efficacy scale, attitudes, and perceptions subject survey instruments. Data from the skill self-efficacy scale, attitudes, and perceptions instruments were summed for subjects assigned to the same team in order to compare subject survey data with team level data provided by each faculty member. Scores on the TPOT subscales and overall scores, behavioral markers, and the NLN Checklist were summed across the three raters who scored each team. These summary scores were added to the individual data set in order to analyze differences in groups based on demographic variables.

DEMOGRAPHICS. The demographic questionnaire consisted of five items: gender, ethnicity, education level, healthcare experience, and experience with team training. Chi square tests of independence were calculated on gender, education, and healthcare experience since these variables are nominal and dichotomous. The analysis was conducted during preliminary data analysis in order to determine if there were any differences across the study variables. In addition, a one-way analysis of variance (ANOVA) was conducted on demographic variables with three or more levels. Ethnicity and Experience with Team Training variables were

compared with dependent variables in the study (perceptions raw score and subscale scores, attitudes raw score and subscale scores, skill self-efficacy raw score, TPOT raw and subscale scores, behavioral markers raw scores, and the NLN Checklist raw scores).

Reliability of Instruments. Cronbach's alpha coefficient was calculated for each instrument used in the study. The individual level data set containing survey data at the student level was used to calculate Cronbach's alpha for the perceptions, attitudes, and skill self-efficacy instruments. Cronbach's alpha for each of the observational instruments (TPOT, behavioral markers, and NLN Checklist) was calculated using the team level data set which consisted of scores for the instruments provided by each rater. Cronbach's alpha was calculated on each instrument by rater as well as for all raters combined.

Interrater Agreement. Intraclass Correlation Coefficients (ICC) were used to calculate interrater agreement for each pair of faculty who shared team ratings. According to Graham et al. (2012) interrater agreement is the "degree to which two or more evaluators using the same rating scale give the same rating to an identical observable situation" (p. 5). Using this definition, it is the absolute value of the rating that is compared. ICC is more meaningful than kappa when there are more than five rating categories or when ratings use a continuous scale (Graham et al., 2012). Since the behavioral markers use a nominal scale (1=present vs. 0=not present) and the TPOT uses a Likert (ordinal) scale for each item, total raw scores were used for each scale to obtain interval level data in order to use the same interrater agreement calculation methodology. Otherwise, various versions of kappa could be used to assess interrater agreement for the nominal data (behavioral markers) and Intraclass Correlation Coefficient could be used to assess agreement on each of the ordinal items on the TPOT.

A one-way random effects analysis of variance with absolute agreement ICC was calculated using total raw scores for pairs of raters for each team. Although each team was rated by three faculty members, it was not always the same faculty members performing the rating, therefore, the one-way model was used (Hallgren, 2012). Absolute agreement was selected because the agreement of the overall ratings was of greatest concern (Hallgren, 2012) rather than rater consistency on each item. Intraclass Correlation focuses on the variance in scores due to performance rather than rater variations, however, results can be misinterpreted if there is limited variation between those being scored (Graham et al., 2012). Each rater provided scores for a select grouping of teams. The method in which raters were assigned to teams is an important design element to consider during the analysis phase. Although the method in which raters were grouped to score teams may not affect the ICC calculations, it does affect the ability to generalize the study results (Hallgren, 2012) and compare TeamSTEPPs® Master Trainer ratings from ratings provided by course faculty. According to Cicchetti (1994) the following cutoff values are recommended for ICC: <.40 = poor, .40 - .59 = fair, .60 - .74 = good, .75 – 1.0 = excellent. A negative ICC indicates substantial disagreement.

Intraclass Correlations provide a single measures and average measures statistic (Hallgren, 2012). Fully crossed study designs/average measures require that ratings for each subject/group are provided by every rater. The ICC is based on the average of the multiple raters. Alternately, in single measures designs, a small subset of the ratings are provided by a sample group of raters, then the remaining ratings are based on scores provided by one rater. Since the study utilized a total of eight raters and eighteen teams, it was not feasible for every rater to provide scores for every team. All teams were rated by three raters, but it was not always the same three raters. The design is not a pure crossed or single measures design which causes some

difficulty in analyzing the results using single or average measures alone. Therefore, both single and average measures will be reported in the results of preliminary analysis. Hallgren (2012) notes that average measures ICC tend to be more reliable than single measures ICC and that when single measures are low and average measures are high, the researcher should report both results.

PRIMARY STUDY ANALYSES

The primary study analyses were conducted after all preliminary analyses were complete in order to ascertain the necessity of data driven changes in the planned analyses for study aims and research questions. Results of the preliminary analysis not only can affect the generalizability of the primary analyses, but can also identify modifications in planned analyses that resulted from unexpected distributions or limitations in the data collected. Implications of the preliminary analyses and the impact on the primary analyses will be discussed in Chapter Four.

AIM 1 HYPOTHESIS 1. In order to address the hypothesis that interrater reliability will be higher for the behavioral markers than the TPOT, one-way random effects model ICCs were calculated. A comparison of average and single measures ICCs was completed during preliminary analyses, therefore, average ICCs for the TPOT and behavioral markers were compared to address the hypothesis. Cicchetti's (1994) cutoff values for ICC were used: <.40 = poor, .40 - .59 = fair, .60 - .74 = good, .75 – 1.0 = excellent. In addition, negative ICC values were indicators of substantial disagreement (Hallgren, 2012).

AIM 1 HYPOTHESIS 2. A within analysis of variance was conducted between the TPOT and behavioral markers in order to address the second hypothesis and determine whether there was a significant difference in the two TPOT assessments. A repeated measures single group

ANOVA was conducted between the TPOT and behavioral markers instrument. The team level database was used to analyze data for the hypothesis. Summary scores from each rater were totaled to create total TPOT and behavioral marker scores for each team. Therefore, each team's score on the TPOT and behavioral marker instrument was a composite score from three raters on each instrument. A single-factor repeated measures design usually compares results from treatment conditions within the same subject and between different subjects (Portney & Watkins, 2009). In this study, differences between the TPOT and behavioral markers are the within subjects comparison. Results from each of the eighteen teams are compared for the between subjects comparisons. Significance was set at the .05 level.

AIM 1 HYPOTHESIS 3. Pearson Product-Moment Correlation Coefficients were calculated between the TPOT, behavioral markers, TeamSTEPPs® attitudes (T-TAQ) and perceptions (T-TPQ). The team data file was used for the analysis. Correlations were calculated using overall raw scores for each instrument and subscale raw scores for the behavioral markers, perceptions, and attitudes. Since the perceptions and attitudes data were originally collected at the individual level, team raw scores were calculated in SPSS by summing scores from the three members on each team for the subscales and overall instrument.

AIM 2 RESEARCH QUESTION 1. A Multitrait-Multimethod Correlation Matrix was created based on literature indicating that the attitudes, perceptions, and original TPOT all measure the construct of teamwork (AHRQ, 2014a; AHRQ, 2014b, AHRQ, 2014c) and therefore, they should theoretically correlate at least moderately. However, no studies were identified which examined this potential relationship. In order to address this research question, Pearson Product-Moment Correlations were calculated between the behavioral markers, original TPOT, perceptions, attitudes, and NLN Simulation Checklist and then entered in to the MTMM

Correlation Matrix. Raw overall scores and raw subscale scores for the TPOT, perceptions, and attitudes were correlated with the overall TPOT and NLN Checklist raw total scores. Moderate (r = .4 - .6) to high correlations (r = .6 - .9) are considered acceptable values to support convergence and provide evidence for validity (Portney & Watkins, 2009, Trochim, 2006).

AIM 2 RESEARCH QUESTION 2. Pearson Product-Moment Correlations were calculated between overall raw scores and subscale raw scores for the TPOT, raw scores for the behavioral markers, and the overall raw team score on the Clinical Skills Self Efficacy Scale. Portney and Watkins (2009) indicate that discriminant/divergent validity is supported when a construct demonstrates low correlations with measurements of a different construct. Correlations of less than .4 between team performance measures and self-efficacy are considered low and indicate divergence.

ETHICAL CONSIDERATIONS

Data were coded to protect the confidentiality of the subjects. Subject survey data submitted through Blackboard were anonymous, therefore subjects were asked to provide their individual, team, and faculty evaluator code numbers in order to merge the data. Paper and pencil evaluations submitted by faculty were coded with team and faculty code numbers. No names were identified on any of the survey forms. The only link to each person's identity was through a code book which was maintained in a locked cabinet in the PI's locked office.

Subjects in the Preceptorship course were provided an option to participate or not participate in the research study. Participation involved completing surveys on Blackboard and receiving faculty teamwork evaluations. However, the teamwork simulation was required for all subjects in the Preceptorship class because it would be unethical to withhold this teamwork experience for subjects who are soon to enter the nursing profession where teamwork is an

expectation. Subjects completing the simulation received a grade of satisfactory. Subjects received no penalty or benefit to their grade by participating in the research study. The student on maternity leave viewed the teamwork education via PowerPoint with voiceover, viewed two simulation videos, and rated two teams using the TPOT, behavioral marker instrument, and the NLN Checklist. The student then answered the same debriefing questions as subjects who participated in real time. Faculty conducted a 1:1 debriefing session with the student to allow the student an opportunity to reflect and review observations.

SUMMARY OF CHAPTER THREE

Chapter Three discussed the study's research design. The study aims, hypotheses, and research questions were outlined along with data management and data analysis strategies. The study protocol which included participant education, instruments, and rater training were also discussed. Ethical considerations were also described.

PLAN FOR REMAINING CHAPTERS

The study findings will be presented in Chapter Four. Implications of the current study design will be described. Discussion, implications of the research findings in this study, areas of future research, and conclusions will be presented in Chapter Five.

Chapter 4 Results

INTRODUCTION

Chapter 4 begins with a basic description of the study sample and overview of the study findings. Preliminary analyses, including demographic comparisons, instrument reliability and interrater agreement are discussed as a precursor to the primary study analyses. Findings from the primary analyses will be discussed according to the research aims, hypotheses, and research questions.

SAMPLE DESCRIPTION

The study sample included 54 senior level baccalaureate nursing students enrolled in the Preceptorship course, a capstone clinical course in the final semester of the nursing program at the University. Subjects were randomly assigned to 18 teams consisting of three team members each. Gender and Ethnicity demographics were completed for all subjects; however, eleven subjects (20.4%) elected not to complete the three remaining demographic questions about education level, healthcare experience, and experience with team training. The following descriptive statistics are based on a total sample size of 54 subjects with the exception of those three variables (n=43).

The sample included African-American/Black (n=10), Asian (n =9 which included subjects identifying as Southeast Asian/Indian), Caucasian (n =30), and Hispanic (n =5) subjects. Thirty-eight subjects (70.4%) out of the total sample (n =54) indicated they were completing their first degree and five subjects (9.3%) indicated they had a previous degree. Thirty subjects (55.6%) indicated they had no healthcare experience prior to nursing school while thirteen subjects (24.1%) indicated they did have some experience. In the explanation section of the survey, subjects stated their healthcare experience included working as nursing assistants,

veterinary assistants, and medical office assistants. Twelve subjects (22.2%) indicated they had no experience with team training, twenty-two subjects (40.7%) indicated they had minimal exposure to team training prior to nursing school, and nine subjects (16.7%) indicated they had formal team training in their work setting.

DESCRIPTIVE ANALYSIS

In this section, raw scores, mean scores, and percentages will be used to provide a basic overview of the data. The preliminary and primary analyses used the raw overall scores for comparison in SPSS. However, since each instrument has a different overall numerical score, means and percentages for each instrument and subscale were computed for descriptive analyses in order to characterize the sample and to provide greater understanding of the comparisons which will be presented later in the Chapter.

Overall scores on the observational instruments used by faculty raters (TPOT, behavioral markers, NLN Checklist) were created by summing ratings across items within each faculty rater and then across all the three faculty raters for each team. These overall scores were then divided by the total number of items on each instrument and then by three (raters) to give an overall mean score on the original TPOT (1=Poor to 5=Excellent). However, each item on the behavioral marker instrument and the NLN Checklist were scored as complete or incomplete. Therefore, mean scores for the behavioral markers and the NLN Checklist are not meaningful. Instead, a percentage score of items completed was calculated for the behavioral markers and the NLN Checklist. Table 1 presents the overall raw scores and mean scores and/or percent of the total possible for each observational instrument. The total possible overall score for the original TPOT was 345, behavioral markers was 51, and the NLN Checklist was 75. Interestingly, the teams who scored highest on one instrument did not necessarily score highest on all measures.

For instance, Team 14 has the second highest TPOT overall score, but the highest behavioral markers (TBM) score and NLN Checklist score. Conversely, team six has the highest TPOT overall score and a substantially lower NLN Checklist percentage score. Differences in rater scores for the observational instruments will be discussed later in the Preliminary Analysis. The full Summary of Observational Instrument Team Scores can be found in Appendix J.

Table 1 Excerpt from Summary of Observational Instrument Team Scores Table (Appendix J)

TEAM	**TPOT Overall**	**TPOT Mean**	**TBM Overall**	**TBM Percent**	**NLN Overall**	**NLN Percent**
1	161	2.36	23	45.10	41	52.56
6	283	4.10	42	82.35	49	62.82
14	278	4.03	44	86.27	71	91.03
16	163	2.36	25	49.02	34	43.59

Survey scores provided by each subject were summed according to team affiliation to create overall team raw scores for the perceptions, attitudes, and skill self-efficacy. The maximum overall score for teamwork perceptions was 525, while the maximum teamwork attitudes score was 450 and the skill self-efficacy maximum was 270. Students rated each individual item on the perceptions and attitudes from one (Strongly Disagree) to five (Strongly Agree). The overall teamwork perceptions, attitudes, and skill self-efficacy scores were divided by the total number of items and then by three (individual subjects) to provide overall mean scores. The mean scores were not used for analyses addressing study questions, but can provide insight into the sample. Students rated each of the 9 items on the CSES on a scale of zero (No Confidence) to ten (Total Confidence). Therefore, mean scores range from zero to ten (see Table 2). In the Table 2 data subset, Team 16 has the highest teamwork perceptions and skill self-

efficacy mean and the second highest teamwork attitudes mean. Conversely, Team 1 has the second highest teamwork perceptions mean, highest teamwork attitudes mean, and lowest skill self-efficacy mean. Interestingly, Team 1 had the lowest TPOT and behavioral marker scores (see Table 1) illustrating an apparent lack of consistency between observed scores (TPOT and behavioral markers) and student attitudes (T-TAQ) and perceptions (T-TPQ) of teamwork. These findings will be addressed in greater detail in the Preliminary Analysis. The full Summary of Overall and Mean Survey Scores by Team can be found in Appendix K. A summary table of all observational and survey scores is available in Appendix L.

Table 2 Excerpt from Summary of Overall and Mean Survey Scores by Team Table

TEAM	T-TPQ Overall	T-TPQ Item Mean	T-TAQ Overall	T-TAQ Mean	CSES Overall	CSES Mean
1	500	4.76	426	4.73	177	6.56
6	490	4.67	378	4.20	228	8.44
14	419	3.99	387	4.30	215	7.96
16	512	4.88	394	4.38	231	8.56

PRELIMINARY ANALYSIS

Preliminary analyses were conducted prior to the primary study analyses to determine whether significant differences existed between groups that might affect interpretation of primary results. Cronbach's alpha was calculated for all six study instruments as an estimate of reliability. In addition, interrater agreement was calculated using intraclass correlations.

DEMOGRAPHIC COMPARISONS

There were no significant differences between males and females on the distribution across education or healthcare experience. A one-way Analysis of Variance (ANOVA) was

calculated comparing ethnic groups on each of the study variables measured by the teamwork perceptions, attitudes, skill self-efficacy, NLN Checklist, TPOT, and behavioral markers. In order to complete the analysis, team scores for the TPOT, behavioral markers, and NLN Checklist were added to the individual level data file. Each member of the team was assigned the team score in order to compare with their individual demographic data. No significant differences were observed between ethnic groups on any of the study variables.

A one-way ANOVA was calculated comparing individual teamwork perceptions and attitudes total scores and their respective subscale scores among subjects with no team training, minimal exposure to team training prior to nursing school, and subjects receiving formal team training in the work setting (see Table 3). The subjects in these three groups did not differ significantly on teamwork perceptions (T-TPQ) total scores or attitudes (T-TAQ) total scores. Based on the standard deviations, there was a greater variability in scores on teamwork perceptions than teamwork attitudes. However, the total possible score on teamwork perceptions is greater than teamwork attitudes, so the greater range of possible scores likely contributes to the variability. There were also no significant differences on the teamwork perceptions *Team Structure*, *Leadership*, or *Communication* subscales; however, significant results were observed on the teamwork perceptions *Situation Monitoring* subscale and the teamwork perceptions TPQ *Mutual Support* subscale. Tukey HSD post hoc analysis for *Communication* and *Situation monitoring* subscales revealed the significant difference was between the group with minimal exposure to team training and the group with formal classes in team training. The group (n = 9) with formal classes in team training had a significantly higher teamwork perceptions *Situation Monitoring* subscale mean scores (32.56) than the group (n = 22) with minimal exposure to team training (mean = 28.95). The group with formal classes in

team training also had significantly higher mean scores on the teamwork perceptions *Mutual Support* subscale (32.56) than the group with minimal exposure to team training (28.77). No significant differences were observed on the teamwork attitudes subscales or the skill self-efficacy total scores. Table 3 depicts descriptive information for the ANOVA comparing groups with varying levels of team training on the study variables.

A one-way Analysis of Variance was also calculated comparing team TPOT total scores, TPOT subscale scores, behavioral markers, and NLN checklist scores between groups with three levels of team training experience (none, minimal, formal training). A significant difference was observed on the TPOT *Communication* subscale score ($p = .025$), *Leadership* subscale ($p = .049$), and *Situation Monitoring* subscale ($p = .032$). Tukey HSD post hoc analysis revealed that the significant difference was between the group with minimal exposure to team training and the group with formal classes in the work setting. However in a surprising finding, the mean scores indicate that the group with minimal team training had higher scores on the TPOT overall and subscale scores than the group who reported to have formal team training prior to nursing school. TPOT *Communication* subscale scores were higher for the minimal team training group ($m = 41.18$, $sd = 6.84$) than the formal team training group ($m = 33$, $sd = 10.15$). Scores for the minimal exposure group ($m = 62.41$, $sd = 10.97$) were significantly higher than the formal team training group ($m = 51.33$, $sd = 15.23$) on the TPOT *Leadership* subscale with greater variation in the formal team training group. Additionally, the TPOT *Situation Monitoring* subscale scores were also higher for the minimal exposure group ($m = 54.09$, $sd = 7.82$) than the formal team training group ($m = 45.11$, $sd = 11.21$). A significant difference was also observed on the behavioral marker instrument ($p = .019$) and the NLN checklist instrument ($p = .012$) between the groups with varying levels of team training. Tukey

HSD post hoc analysis revealed the difference was between the group with minimal exposure to team training prior to nursing school and the group with formal classes in the work setting. Similar to the findings from the TPOT instrument, scores for the group reporting formal team training (TBM mean = 30.56, sd= 8.69; NLN mean = 44.44, sd = 8.08) were significantly lower on the behavioral marker instrument than scores for students reporting minimal team training (TBM mean = 37.91, sd = 6.02; NLN mean = 53.36, sd = 7.27). These are surprising results; however, it is important to note that subjects report receiving formal team training but the amount or quality of training was not quantified. Table 3 displays ANOVA results for Experience with Team Training on all the total instrument scores and only the subscale scores which were significant. Since eleven students did not answer the demographic question about experience with team training, the total n = 43.

Table 3 ANOVA Results for Experience with Team Training (n=43)

Instrument	**Subscale**	**Df**	**F**	**Mean**	**SD**	**Significance**
TPQ	Overall	2,40	3.145	151.23	15.78	.054
	Situation Monitoring	2,40	3.652	29.95	3.58	.035*
	Mutual Support	2,40	4.435	30.12	3.59	.018*
TAQ	Overall	2.40	2.444	128.63	9.33	.100
CSES	Overall	2,40	.381	72.79	9.27	.686
TPOT	Overall	2,40	2.894	228.21	38.87	.067
TPOT	Communication	2,40	4.037	38.67	7.80	.025*
	Leadership	2,40	3.262	60	12.05	.049*
	Situation Monitoring	2,40	3.760	51.93	8.99	.032*
TBM	Overall	2,40	4.369	36.16	6.93	.019*
NLN Checklist	Overall	2,40	4.968	50.77	7.80	.012*

**Note: Tukey HSD post-hoc analysis revealed the significant difference was between the group with minimal exposure to team training and the group with formal classes in team training ($p < .05$).*

RELIABILITY OF INSTRUMENTS

Cronbach's coefficient alpha was calculated to examine internal consistency of the teamwork perceptions, attitudes, TPOT, behavioral markers, skill self-efficacy, and NLN Checklist. Subjects (n = 54) individually completed the teamwork perceptions, attitudes, and skill self-efficacy surveys; however, some missing data was observed. Team subscale and overall survey instrument scores were summed without the missing data. The teamwork perceptions survey (T-TPQ) consists of 35 items and is measured on a five point Likert-type scale. Cronbach's alpha for this sample was .977 (n = 43). The teamwork attitudes (T-TAQ) is a 30 item instrument which measures teamwork attitudes on a five point scale. Cronbach's alpha for the current study was .901 (n = 48). The CSES is a nine item instrument measuring subjects' self-efficacy with clinical skills on an eleven point scale (0 = no confidence to 10 = total confidence). Cronbach's alpha for this instrument was .843 (n = 53).

A total of eight faculty rated groups on the TPOT, behavioral markers, and the NLN Checklist. No missing data was observed. The TPOT is a 23 item instrument on which raters provide scores from one (very poor) to five (excellent) for each item. The behavioral markers instrument is a 13 item instrument on which raters determine whether the specific teamwork behavioral markers are present or absent. The NLN Checklist is a 26 item instrument on which essential simulation interventions are scored as present or absent. Table 4 displays Cronbach's alpha for the TPOT, behavioral markers, and the NLN Checklist according to instrument scores provided by the individual TeamSTEPPs Master Trainers (raters 6, 7, 8) and the Cronbach's alpha for each instrument with ratings from all raters combined. Reliability coefficients >.75 are considered good according to Portney and Watkins (2009). When all raters are combined, the internal consistency improves; however, with increased numbers, increased alpha coefficients are

expected. It is important to note that consistency on instruments at the individual faculty level falls below .75 for the behavioral marker instrument and the NLN Checklist in several cases.

Table 4 Cronbach's Alpha by TeamSTEPPs® Master Trainer Raters and All 8 Raters Combined

RATER	N groups evaluated	TPOT	TBM	NLN Checklist
6	9	.871	.490	.711
7	18	.975	.775	.573
8	9	.987	.526	.627
All Raters (1,2,3,4,5,6,7,8)	18	.979	.785	.747

INTERRATER AGREEMENT

Analysis of interrater agreement is included in the preliminary analysis because the primary analyses are contingent upon reliable observational scores. A total of eight faculty observed and scored teams of three nursing subjects on the TPOT, behavioral markers, and the NLN Checklist. Three faculty are TeamSTEPPs® Master Trainers which means they attended a three day teamwork workshop developed by the Agency for Healthcare Research and Quality in May of 2015. The other five faculty members participated in a five and one half hour workshop on teamwork, utilization of the observational instruments, and interrater reliability with the observational instruments used in the study. In the study, eighteen teams were evaluated which resulted in a small number of rater comparisons. Intraclass correlations can be affected by variability among raters as well as variability within the same rater's scores across teams; therefore, single and average measures ICCs were reported.

Two of the TeamSTEPPs® Master Trainers have less than five years teaching experience and minimal simulation experience. The third TeamSTEPPs® Master Trainer, who is also the PI for the current research study, has more than 12 years teaching experience and 11 years simulation experience. The remaining five faculty members had from one to fifteen years teaching experience and from one to five years simulation experience. With the exception of the PI, only three faculty consider themselves to be very experienced with simulation, two have moderate experience, and two have minimal experience. Since faculty teaching the simulations did not receive education directly from TeamSTEPPs®, the PI elected to ask two of the TeamSTEPPs® Master Trainers to each rate half of the teams so that each team was evaluated by two experienced TeamSTEPPs® Master Trainers. The PI rated all teams in order to have one person consistently measuring all groups. Therefore, the design in the current study does not completely meet the definitions required for using either the single measures ICC or average measures ICC as described in Chapter 3. The design in this study is somewhat of a hybrid which was a pragmatic design decision since only three faculty at the University are TeamSTEPPs® Master Trainers and faculty who were available to participate in the study have varied levels of simulation experience.

Each team of students (n=18) was rated by three faculty members: the PI, an additional TeamSTEPPs® Master Trainer, and a faculty member. Table 5, Intraclass Correlation Coefficient (ICC) Comparison for Interrater Agreement, compares single and average ICCs for each of the rater pairings.

Table 5 Intraclass Correlation Coefficient (ICC) Comparison for Interrater Agreement

	TBM		TPOT	
RATERS	Single Measures	Average Measures	Single Measures	Average Measures
6 & 7	.501	.667	.481	.649
7 & 8	.504	.670	.566	.723
3 & 7	.625	.769	.639	.780
5 & 7	.241	.389	.913	.954
3 & 8	.518	.682	.630	.773
5 & 8	.376	.547	.625	.770
4 & 7	.458	.628	.889	.942
4 & 6	.436	.608	.473	.642
1 & 7	.126	.224	.340	.508
1 & 6	-.429	-1.5	-.296	-.840
2 & 7	.319	.484	.227	.370
2 & 6	-.667	-4.0	-.529	-2.25
2 & 8	-.308	-.889	.797	.887
1 & 8	Unable to calculate			
5 & 6	Unable to calculate			

The ICCs were calculated using single random measures with absolute agreement because it is the overall agreement in score that is most important for the current study. Table 5 shows that the single and average ICC measures exhibited some differences for the rater pairings. Given the greater stability of average measures (Hallgren, 2012) and the reflection of systematic rating differences of multiple raters as in this study, in subsequent discussion, average measures will be the focus of interest.

Table 5 illustrates that TeamSTEPPs® Master Trainers, Raters 6, 7, and 8 have similar single and average measures ICCs. When scoring teamwork with the behavioral markers (TBM) instrument, raters 6 and 7 have an average measures ICC of .667 and raters 7 and 8 have average measures ICC .670. Both of these ICCs are considered in the "good" range according to Cicchetti (1994). The ICC for raters 6 and 7 is lower (.649) for the TPOT while the ICC for raters 7 and 8 is higher (.723). Overall, the TeamSTEPPs® Master Trainers demonstrated good

interrater agreement. Rater 4 shared ratings with TeamSTEPPs® Master Trainers 6 and 7. The ICCs between raters 4 and 7 and between raters 4 and 6 were very close and considered in the fair range. Similarly, TeamSTEPPs® Master Trainers 7 and 8 both shared ratings with raters 3 and 5. The average measures ICCs between raters 3 and 7 were considered good and also close to the ICCs between raters 3 and 8. However, average measures ICCs for raters 5 and 7 were poor on the behavioral markers instrument (TBMs) while average measures between raters 5 and 8 were fair. Raters 1 and 2 demonstrated fair to poor ICCs with TeamSTEPPs® Master Trainers 6, 7, and 8 on behavioral marker ratings. When reviewing ICCs on the TPOT for raters 1, 2, 6, 7, and 8, only the pairing of raters 2 and 8 demonstrated acceptable ICCs. However, raters 2 and 8 demonstrated negative average measure ICCs on the behavioral markers instrument which indicates systematic disagreement according to Hallgren (2012). Pairings of raters 1 and 8 as well as raters 5 and 6 could not be calculated because these raters only shared one group. Table 6 illustrates the simple percentage of agreement on the behavioral marker items for rater pairs 1 and 8 as well as raters 5 and 6 since an ICC could not be calculated. Since the raters were neither completely random nor grouped uniformly to rate the same teams, the correlation matrix was unstable (Pedhazur & Schmelkin, 1991) and resulted in two values outside of the expected -1 to 1 (see values for raters 1 & 6 and for raters 2 & 6 depicted in Table 5). Future studies using multiple raters need to carefully consider rater groupings in order to prevent this error. Results of these interrater agreement (ICC) analyses should be interpreted with caution because the design was not complete and did not meet all the requirements for using single or average measures ICCs alone. Due to these results, further analyses of the hypotheses and research questions should also be interpreted with caution.

Table 6 Agreement Matrix for Rater Pairings with No ICC

	Raters								
TPOT with TBMs		**5**	**6**	**Agree**	**Disagree**	**1**	**8**	**Agree**	**Disagree**
	1A	1	1	X		1	0		X
	1B	1	1	X		1	1	X	
	2A	1	1	X		0	0	X	
	2B	1	0		X	1	1	X	
	3A	1	1	X		1	1	X	
	3B	1	0	X		1	1	X	
	3C	1	0		X	1	1	X	
	4A	1	1	X		1	1	X	
	5A	1	1	X		1	1	X	
	6A	1	1	X		1	1	X	
	7A	1	1	X		1	1	X	
	8A	1	1	X		1	0		X
	9A	0	1		X	0	0	X	
	10A	1	0		X	1	1	X	
	11A	1	1	X		1	1	X	
	12A	1	1	X		1	1	X	
	12B	1	1	X		1	1	X	
				76.47% Agreement				88.24% agreement	

PRIMARY ANALYSIS

AIM 1 HYPOTHESIS 1

The first hypothesis asserts that interrater agreement on simulation performance will be higher for behavioral markers (TBMs) than the original TPOT demonstrating higher rating reliability and specificity. In order to address the hypothesis, one-way random effects ICCs with absolute agreement were calculated between raters who evaluated the same teams (see Table 5). Interrater agreement was assessed for a total of fifteen rater pairs. When comparing ICCs for the TPOT and behavioral markers, only the pairing between raters 6 and 7 (6.67%) demonstrated greater ICC on the behavioral marker instrument (average measures = .667). Average measures ICC for the behavioral markers was also greater than the TPOT for raters 2 and 7 (6.67%); however, both values were poor. Nine rater pairings (60%) demonstrated ICCs which were

greater on the TPOT than the behavioral markers. Three rater pairings (20%) demonstrated negative ICCs on the behavioral markers and two rater pairings (13.33%) revealed negative ICCs on the TPOT. Negative average measures ICCs indicate systematic disagreement between raters on the absolute score (Hallgren, 2012). For two of the rater pairings (13.33%), ICCs were unable to be calculated because the raters only shared one group (see Table 6).

The ICC values presented do not support Aim 1, Hypothesis 1. However, ratings were comparable on the TPOT and behavioral markers between the TeamSTEPPs® Master Trainers which is consistent with the framework asserted in the Teamwork Nomological Network. Several factors may have influenced these results: faculty expertise in TeamSTEPPs® Teamwork Education, faculty proficiency in evaluating simulations, the level of difficulty in evaluating each team's performance, the order in which evaluation instruments were completed, and the reliability of the newly developed behavioral markers instrument.

Each team was rated by two TeamSTEPPs® Master Trainers and one faculty member. Therefore, interrater agreement comparisons between TeamSTEPPs® Master Trainers and faculty members were not conducted. At the University, faculty use simulation largely as a teaching strategy; therefore, rigorous evaluation of simulation experiences is not routinely conducted. Simulation rating for performance evaluation may be variable between faculty members. In addition, each team performs differently during simulation. Some teams may be more difficult to rate, thereby creating additional variability amongst raters. Faculty anecdotally expressed the greatest difficulty completing the TPOT instrument because it lists generic statements and asks the rater to score the group on a scale from 1 to 5 for each statement. Every faculty completed the NLN Checklist and behavioral markers during the simulation and saved the TPOT to complete after the simulation. Therefore, since all faculty completed the

behavioral markers prior to the TPOT, their answers on the TPOT may have been influenced by exposure to the behavioral markers. It is plausible that raters provided more consistent ratings on the TPOT due to their exposure to the behavioral markers. However, since all raters utilized both observational instruments (TPOT and behavioral markers), the potential effect from exposure to the behavioral markers cannot be analyzed in this study. Lastly, the behavioral markers instrument is new and lacks established reliability data. It is plausible that faculty interpreted and scored the instrument differently and inconsistently due in part to variability in each team's performance as well as potential ambiguity of some of the behavioral markers items. As discussed in Chapter 3 and the Preliminary Analysis section of this Chapter, results of these analyses should be interpreted with caution since this study's research design did not fully meet the requirements for utilizing single or average measures ICCs alone and the reasons for variability in ratings is not clear.

AIM 1 HYPOTHESIS 2

The second hypothesis proposes there will be a significant difference in behavioral marker (TBM) and TPOT assessments with the behavioral markers demonstrating better performance assessment. A within analysis of variance was conducted between the TPOT and behavioral markers in order to address the second hypothesis and determine whether there is a significant difference in the two assessments. A one-way repeated measures ANOVA was calculated comparing results on the TPOT and behavioral markers for all eighteen teams. Since differences between the TPOT and behavioral markers are the within subjects comparison, it was necessary to transform both variables to a common metric, z scores, for the purpose of analyses. No significant effect was found ($F(1,17) = .000$, $p = 1.000$) between the TPOT and behavioral markers instruments indicating there is no difference in the scores.

AIM 1 HYPOTHESIS 3

The third hypothesis proposes that correlations between the TPOT, behavioral markers (TBMs), teamwork attitudes and perceptions will be higher for the behavioral markers than the original TPOT demonstrating greater sensitivity. Pearson Product-Moment Correlation Coefficients were calculated between the TPOT, behavioral markers, teamwork attitudes, and perceptions surveys. The team data file was used to calculate correlations between the TPOT, behavioral markers, teamwork perceptions and attitudes. Raw scores for each instrument as well as subscale raw scores for the TPOT, behavioral markers, teamwork perceptions, and attitudes were correlated (see Table 7). The TPOT and behavioral markers similarly demonstrated weak to fair correlations with teamwork perceptions and attitudes overall scores which indicates there is not a strong relationship between observed teamwork behavior, perceptions, and attitudes in this study sample as measured by the instruments overall. The relationship between behavioral markers and the total teamwork perceptions score was slightly stronger (r = -.354) than the relationship between the TPOT and total teamwork perceptions (r = -.315), yet both were negative. The correlation between behavioral markers and total attitudes score was also slightly stronger (r = .191) than the correlation between the TPOT and the total attitudes score (r = .028); however, both demonstrate little to no relationship.

The exceptions were the teamwork perceptions *Situation Monitoring* subscale which demonstrated moderate negative correlations with behavioral markers (r = -.520), TPOT (r = -.516), and TPOT subscales of *Communication (r = - .498), Leadership (r = - .551), and Situation Monitoring (r = - .520).* Moderate negative correlations were also observed between the teamwork perceptions *Communication* subscale and the behavioral markers and the TPOT

Situation Monitoring subscale. The teamwork perceptions *Leadership* subscale also demonstrated moderate negative correlations with the TPOT Overall score, *Team Structure* subscale, *Situation Monitoring* subscale, and the *Mutual Support* subscale. Only correlations between the teamwork perceptions *Situation Monitoring* subscale, behavioral markers, TPOT Overall score, and the TPOT *Communication, Leadership, and Situation Monitoring* subscales reached the level of significance ($p < .05$). These findings indicate that there is a moderate yet negative relationship between these variables. Table 7 displays the correlations between the TPOT, behavioral markers, perceptions and attitudes overall and subscales. In all but three instances (teamwork attitudes *Leadership*, *Situation Monitoring* and *Communication* subscales), correlations with behavioral markers were slightly higher than correlations with the original TPOT; therefore, Aim 3, Hypothesis 3 is in part supported.

Table 7 Correlations between TPOT, Behavioral Markers (TBM), Teamwork Perceptions (T-TPQ) and Teamwork Attitudes (T-TAQ) Total and Subscale Raw Scores

	T-TPQ Overall	TPQ-TS Subscale	TPQ-L Subscale	TPQ-SM Subscale	TPQ-MS Subscale	TPQ-C Subscale	T-TAQ Overall	TAQ-TS Subscale	TAQ-L Subscale	TAQ-SM Subscale	TAQ-MS Subscale	TAQ-C Subscale
TBM	-.354	-.166	-.221	-.520*	-.243	-.435	.191	-.184	-.340	-.259	-.153	-.038
TPOT Overall	-.315	-.151	-.146	-.516*	-.226	-.365	.028	-.165	-.412	-.274	-.091	-.043
TPOT (TS Subscale)	-.156	-.068	-.035	-.315	-.054	-.214	-.024	-.131	-.440	-.303	.098	-.177
TPOT (C Subscale)	-.346	-.203	-.195	-.498*	-.316	-.338	-.053	-.200	-.338	-.306	-.128	-.037
TPOT (L Subscale)	-.309	-.103	-.088	-.551*	-.231	-.398	.152	-.078	-.298	-.233	-.266	-.007
TPOT (SM Subscale)	-.342	-.188	-.201	-.520*	-.216	-.406	.048	-.196	-.455	-.250	-.004	-.051
TPOT (MS Subscale)	-.260	-.134	-.150	-.427	-.188	-.260	-.091	-.192	-.434	-.198	.009	.037

**Note: Correlations reach the* $p < .05$ *level of significance.*

AIM 2 RESEARCH QUESTION 1

The first research question for Aim 2 examines the relationship between the TPOT, behavioral markers, teamwork perceptions (T-TPQ), teamwork attitudes (T-TAQ), and the NLN Simulation Checklist in order to assess for convergent validity. Pearson Product-Moment Correlations were calculated between raw scores on the TPOT, behavioral markers, teamwork perceptions, attitudes, and the NLN Simulation Checklist (see Table 8). Tables of the complete correlation results can be found in Appendix M and N. As predicted in the Teamwork Nomological Network, a strong positive relationship of large magnitude was found between the TPOT and behavioral markers (r^2 = .87). The NLN Checklist also demonstrated a strong positive correlation with the TPOT (r^2 = .62) and the behavioral markers (r^2 = .63). In addition, the relationship between teamwork perceptions and attitudes overall scores showed a significant positive relationship (r^2 = .40). When teamwork attitudes and perceptions subscales were correlated, significant and positive moderate to excellent correlations were observed between all subscales except for the teamwork attitudes *Mutual Support* subscale (see Table 8). The teamwork attitudes *Mutual Support* subscale has three reverse coded items which may have been confusing to subjects and caused the incongruence. The most significant correlations were observed between the teamwork attitudes *Team Structure, Situation Monitoring,* and *Communication* subscales and the teamwork perceptions subscales; however, significant and positive moderate to good correlations were also observed for the teamwork attitudes *Leadership* subscale and the teamwork perceptions subscales.

Table 8 Correlations Between Teamwork Perceptions (T-TPQ) and Attitudes (T-TAQ)

	T-TPQ Overall	**TPQ-TS**	**TPQ-L**	**TPQ-SM**	**TPQ-MS**	**TPQ-C**
T-TAQ Overall	.633**	.611**	.703**	.435	.588*	.563*
TAQ-TS	.768**	.815**	.799**	.595**	.748**	.555*
TAQ-L	.540*	.480*	.579*	.465	.490*	.450
TAQ-SM	.674**	.640**	.629**	.536*	.684**	.589*
TAQ-MS	.249	.226	.391	.078	.173	.284
TAQ-C	.721**	.694**	.753**	.488*	.706**	.660**

The observational performance evaluations (TPOT, behavioral markers, NLN Checklist) do not strongly correlate with the teamwork perceptions and attitudes overall scores. However, correlations between all variables and the teamwork perceptions *Situation Monitoring* subscale were moderate to good (see Table 9). The teamwork attitudes *Leadership* subscale was the only teamwork attitudes subscale with fair correlations to the other study variables. Therefore, convergent validity was in part established between the TPOT, behavioral markers, NLN Checklist, and the teamwork perceptions *Situation Monitoring* subscale, but not with teamwork attitudes or teamwork perceptions overall scores (see Table 10 which depicts correlations in the Multi-trait Multi-method Correlation Matrix). The finding that overall scores for teamwork perceptions and attitudes do not strongly correlate with observed teamwork performance could be important because teamwork perceptions, obtained through surveys, are often measured in lieu of observed performance in order to measure outcomes of teamwork education. However, due to interrater agreement issues in this study, it is difficult to draw this conclusion with confidence. Further study with larger samples, fully crossed designs for interrater agreement, and reliable and tested observational instruments are needed.

Table 9 Excerpt from the Correlations between TPOT, Behavioral Markers (TBM), Teamwork Perceptions (T-TPQ), and Teamwork Attitudes (T-TAQ)

	TPOT	TPOT (TS Subscale)	TBM	NLN Checklist	T-TPQ	T-TPQ (SM Subscale)	T-TAQ	T-TAQ (L Subscale)
TPOT	1	.777**	.933**	.789**	-.315	-.516*	.028	-.412
TPOT (TS Subscale)	.777**	1	.714**	.541*	-.156	-.315	-.024	-.440
TBM	.933**	.714**	1	.791**	-.354	-.520*	.191	-.340
NLN Checklist	.789**	.541*	.791**	1	-.357	-.490*	-.015	-.325
T-TPQ	-.315	-.156	-.354	-.357	1	.905**	.633**	.540*
T-TPQ (SM Subscale)	-.516*	-.315	-.520*	-.490*	.905**	1	.435	.465
T-TAQ	0.028	-.024	.191	-.015	.633**	.435	1	.891**
T-TAQ (L Subscale)	-.412	-.440	-.340	-.325	.540*	.465	.891**	1

***Correlation is significant at the 0.01 level (2-tailed)*
**Correlation is significant at the 0.05 level (2-tailed)*

AIM 2 RESEARCH QUESTION 2

The second research question for Aim 2 examines the relationship between the TPOT, behavioral markers, and the Clinical Skills Self-efficacy Scale in order to assess for divergent validity. Pearson Product-Moment Correlations were calculated between overall raw scores and

subscale raw scores for the TPOT, raw scores for behavioral markers, and the overall raw team score on the Clinical Skills Self Efficacy Scale (CSES) (see Table 10). Results from this sample indicate that self-efficacy with clinical skills (CSES) is strongly and positively correlated with the TPOT (r^2 = .32) and the behavioral markers (r^2 = .42). This result indicates that self-efficacy with clinical skills is associated with observed teamwork performance; however, the nature of the relationship or cause cannot be determined from this study. Therefore, correlation results between the TPOT, behavioral markers, and the CSES from this sample do not demonstrate divergent validity. It is possible that clinical skill self-efficacy plays a role in teamwork as evidenced by the moderate to strong positive correlations between teamwork behaviors (TPOT and behavioral markers) and self-efficacy with clinical skills (CSES). However, there are interrater agreement issues in this study which make it difficult to draw this conclusion with confidence. Further study with larger samples, reliable and tested observational instruments are needed. The full MTMM Correlation Matrix displaying the correlations described in Aim 2, Research Questions 1 and 2 can be found in Table 11.

Table 10 CSES, TPOT, and Behavioral Markers (TBM) Correlations for Divergent Validity

	TBM	TPOT	TPOT TS Subscale	TPOT C Subscale	TPOT L Subscale	TPOT SM Subscale	TPOT MS Subscale
CSES	.647**	.566*	.497*	.540*	.641**	.437	.456

***Correlation is significant at the 0.01 level (2-tailed)*
**Correlation is significant at the 0.05 level (2-tailed)*

Table 11 Multitrait-Multimethod Correlation Matrix: Teamwork Construct Validity Testing

		TBM	TPOT	T-TPQ	T-TPQ (SM Subscale)	T-TAQ	NLN Checklist	CSES
		Behavior	Behavior	Perception	Perception	Attitudes	Performance	Skill
TBM	Behaviors	1.0						
TPOT	Behaviors	.933** Expected moderate to high	1.0					
T-TPQ	Perception	-.354 Expected moderate to high	-.315 Expected moderate to high	1.0				
T-TPQ (SM Subscale)	Perception	-.520* Expected moderate to high	-.516* Expected moderate to high	.905** Expected moderate to high	1.0			
T-TAQ	Attitudes	.191 Expected moderate to high	-.190 Expected moderate to high	.633** Expected moderate to high	.435 Expected moderate to high	1.0		
NLN Checklist	Performance	.791** Expected moderate to high	.789** Expected moderate to high	-.357 Expected moderate to high	-.490* Expected moderate to high	-.015 Expected moderate to high	1.0	

CSES	Skills	.647** Expected low	.566* Expected low	.125 Expected low	-.061 Expected low	.250 Expected low	.330 Expected low	1.0

SUMMARY

In Chapter 4, findings from the descriptive, preliminary, and primary analyses were presented. The study sample consisted of eighteen teams comprised of a total of 54 individual subjects. The teams were not significantly different based on gender, ethnicity, previous education, or healthcare experience; however, differences were observed when comparing individuals with minimal previous exposure to team training and formal classes in team training. It is not surprising to see differences in teamwork attitudes and perceptions between groups with varying levels of teamwork training. However, it is surprising to see poorer observed teamwork performance in groups who have a history of participating in team training.

The preliminary analysis revealed a difference in interrater agreement. Due to the method in which raters were grouped, it is difficult to determine whether differences in ratings can be attributed to differences in team performance, rater training, ambiguity of the instruments, or whether one instrument mediated rater reliability on another instrument. Since interrater agreement is moderate to good at best, subsequent primary analyses need to be interpreted with caution.

In the primary analysis, interrater agreement for behavioral markers (TBM) was not always better than interrater agreement for the original TPOT. The primary analysis did reveal differences in interrater agreement between the TPOT and behavioral markers. Convergence was in part demonstrated in the correlations between the TPOT and behavioral markers and also between teamwork perceptions and attitudes. However, correlations between the TPOT, behavioral markers, perceptions and attitudes overall scores were poor. One exception was the

teamwork perceptions *Situation Monitoring* subscale which demonstrated positive moderate correlations with observed teamwork behaviors (TPOT and behavioral markers). When the NLN Checklist was added to the correlations, strong positive correlations were observed between the NLN Checklist, TPOT, and behavioral markers, but correlations with teamwork perceptions and attitudes were low. Divergent validity was not established between the TPOT, behavioral markers, TPOT subscales, and clinical skills self-efficacy. These findings, conclusions, implications and limitations of this study will be discussed in greater detail in Chapter 5.

Chapter 5 Conclusions, Discussion, Recommendations

The Chapter will provide an overview of the study purpose, significance, hypotheses and research questions. A description of the sample and major variables will be provided along with a discussion of the study's major findings and conclusions. The Chapter will conclude with a discussion of feasibility, limitations, and recommendations for future research.

Teamwork has been identified as an essential competency by the NAM (IOM, 2003) and QSEN (2012). Despite strong assertions from these national agencies, reports of objective teamwork assessment and evaluation in undergraduate nursing programs are limited. Healthcare providers, particularly nursing students, need to practice teamwork in a safe learning environment and receive specific feedback in order to improve performance (Billings & Halstead, 2012). The opportunity to improve teamwork is critical to student learning and needs to occur prior to entering the high stakes healthcare workforce.

STATEMENT OF THE PROBLEM

Nursing practice is dynamic, necessitating nursing education to evolve to meet the challenge of preparing future nurses to function in this complex healthcare environment. The ability to function effectively in the healthcare team requires practice and feedback in order to improve. Simulation provides an avenue to practice and improve teamwork behaviors in a safe learning environment. A multitude of clinical situations can be replicated through simulation. However, reliable teamwork evaluation instruments are needed in order to provide feedback on team members' performance. In addition, these newly developed instruments need to be tested using a reliable and valid methodology.

OVERVIEW OF THE METHODOLOGY

This methodological study aimed to demonstrate the viability of the Team Performance Observation Tool enhanced with Targeted Behavioral Markers (TBMs) as a sensitive and valid measure of teamwork performance and a superior approach to teamwork assessment without behavioral markers. A Multitrait-Multimethod research design was used with an exemplar convenience sample (n=54) of baccalaureate nursing students to correlate results from the TPOT, behavioral markers, teamwork perceptions (T-TPQ), teamwork attitudes (T-TAQ), and simulation intervention checklist (NLN Checklist) for convergent validity and with the Clinical Skills Self-efficacy Scale (CSES) for divergent validity. Three theoretical frameworks guided the development and implementation of this study: TeamSTEPPs® Educational Framework, the theory of nomological networks, and the Nursing Education Simulation Framework. TeamSTEPPs® is a teamwork instructional framework for healthcare providers. The framework includes educational tools and strategies to teach the identified competencies and instruments for measuring teamwork attitudes, perceptions, and performance. The theory of nomological networks was used to create the Teamwork Nomological Network. When nomological network theory was applied to the concept of teamwork and established teamwork measurement instruments, the Teamwork Nomological Network was created. The Nursing Education Simulation Framework was used to guide the development and implementation of the teamwork simulation experience in this study.

THE STUDY PURPOSE

The purpose of the study was to analyze psychometric properties of the Targeted Behavioral Markers (TBMs) for teamwork using a methodology by which validation of behavioral markers can be efficiently accomplished and replicated with various teamwork

simulation scenarios. The study had two primary aims, three hypotheses, and two research questions.

AIMS, HYPOTHESES, AND RESEARCH QUESTIONS

The first aim of the research was to demonstrate the viability of the behavioral markers (TBMs) as a sensitive and valid measure of teamwork performance assessment in simulation training. Three hypotheses for Aim 1 were tested.

A1H1: Interratter agreement on simulation performance will be higher for the behavioral markers than the original TPOT demonstrating higher rating reliability and specificity.

A1H2: There will be a significant difference in the behavioral marker and TPOT assessments post simulation training demonstrating better performance assessment with behavioral markers.

A1H3: Correlations between teamwork performance evaluations (TPOT and behavioral markers) with TeamSTEPPs® attitudes and perceptions will be higher for the behavioral markers instrument than the original TPOT demonstrating greater sensitivity.

The second aim was to assess the convergent and divergent validity of the Targeted Behavioral Markers (TBMs) as a valid measure of teamwork skill and performance acquisition in simulation training. Two research questions were tested.

A2RQ1: What is the relationship between TPOT, behavioral markers (TBMs), teamwork perceptions (T-TPQ), teamwork attitudes (T-TAQ), and NLN Simulation Checklist (assess convergent validity)?

A2RQ2: What is the relationship between the TPOT, behavioral markers (TBMs), and the Clinical Skills Self-efficacy Scale (assesses divergent validity)?

MAJOR STUDY FINDINGS AND CONCLUSIONS

The following section outlines the major study findings and conclusions. A summary of the study sample and major study variables will be reviewed. This section also describes the preliminary analysis which consisted of demographic comparisons and baseline interrater agreement. Major study findings from the primary analysis include interrater agreement, sensitivity of observatonal instruments, evidence of convergent and divergent validity and conclusions regarding the MTMM Correlation Matrix.

SAMPLE

The study sample consisted of senior level baccalaureate nursing students (n = 54) enrolled in the Preceptorship nursing course, a clinical capstone course in the final semester of the nursing program at the University. Subjects were randomly assigned to 18 teams consisting of three team members each. Gender and ethnicity demographics were completed for all subjects; however, eleven subjects (20.4%) elected not to complete the three remaining demographic questions about education level, healthcare experience, and experience with team training. The following descriptive statistics are based on a total sample size of 54 subjects.

The sample was comprised of African-American/Black (n=10), Asian (n =9) which included subjects identifying as Southeast Asian/Indian, Caucasian (n =30), and Hispanic (n =5) subjects. Thirty-eight subjects (70.4%) out of the total sample (n =54) indicated they were completing their first degree and five subjects (9.3%) indicated they had a previous degree. Thirty subjects (55.6%) indicated they had no healthcare experience prior to nursing school while thirteen subjects (24.1%) indicated they did have some experience. In the explanation

section of the survey, subjects specified that their healthcare experience included working as nursing assistants, veterinary assistants, and medical office assistants. Twelve subjects (22.2%) indicated they had no experience with team training, twenty-two subjects (40.7%) indicated they had minimal exposure to team training prior to nursing school, and nine subjects (16.7%) indicated they had formal team training in their work setting.

MAJOR STUDY VARIABLES

The study examines relationships in this sample between four measures of teamwork and the relationship between these four measures with clinical skill self-efficacy which is a theoretically unrelated concept. Elements of teamwork can be measured through rater observation with the TPOT and behavioral markers. In addition, teamwork may be measured using the teamwork attitudes and perceptions surveys. Therefore, it was hypothesized that all four measures of teamwork would produce similar findings and demonstrate convergent validity. The Clinical Skills Self-Efficacy Scale (CSES) is a self-reported measure of self-efficacy with clinical skills. It was hypothesized that clinical skill self-efficacy is different enough from teamwork to demonstrate divergent validity.

PRELIMINARY ANALYSIS

Preliminary analysis was conducted to determine if there were any differences between subjects based on demographics. In addition, baseline interrater agreement analyses were conducted to determine the reliability of rater observational scores.

DEMOGRAPHIC COMPARISONS. Demographic variables were compared with study variables using Chi Square and ANOVA. Demographic groups were not significantly different based on gender, ethnicity, healthcare experience or previous education level. However, differences were observed in groups with varying levels of team training. The group with

minimal previous exposure to team training demonstrated higher observed performance scores on the behavioral markers instrument, NLN Checklist, and the original TPOT subscales of *Communication, Leadership, and Situation Monitoring.* These findings indicate that the group with minimal teamwork exposure prior to this experience actually communicated better, demonstrated greater leadership skills, and had a greater shared understanding of the situation than the group who stated they received previous formal education in teamwork. This is a surprising finding; however, previous formal teamwork education was not described. It is also possible that the group with minimal previous teamwork education exposure may have valued the teamwork education more and thus tried harder to apply the new education to the teamwork simulation experience. However, due to interrater concerns, these conclusions may not be drawn with confidence.

BASELINE INTERRATER AGREEMENT. Interrater Agreement was calculated for all rater pairs using Intraclass Correlation (ICC). Three TeamSTEPPs® Master Trainers and five faculty raters observed and scored teams of three nursing subjects on the TPOT, behavioral markers, and the NLN Checklist. Two TeamSTEPPs® Master Trainers and one faculty member were assigned to rate each group in order to provide ratings from experienced and inexperienced raters for each group. Average measures for the ICCs were good for the TeamSTEPPs® Master Trainer pairs on the behavioral markers instrument and the TPOT. Ratings (ICC) between faculty members 3, 4, and 5 and the TeamSTEPPs® Master Trainers were similar to the ICC values between the TeamSTEPPs® Master Trainers. Two rater pairings were unable to be calculated because they only shared one group. In addition, the remaining two raters (1 and 2) seemed to be outliers in their ratings demonstrating negative ICC values, which indicate

substantial disagreement. Incidentally, these two faculty members had the least teaching and simulation experience of the group (1-2 years).

In a systematic review of behavioral marker systems in healthcare, Dietz et al. (2015) concluded that researchers need to provide recurrent rater training and establish a standard of reliability reporting. When all rater data was combined in the current study, Cronbach's Alpha was 0.979 for the TPOT which is consistent and slightly greater than findings from Maguire et al. (2014) who demonstrated a Cronbach's Alpha of 0.965 for the TPOT in a sample of thirty-one raters scoring the same simulation. Zhang (2015) used weighted kappa to calculate interrater reliability for one pair of raters using the behavioral marker instrument; however, these two raters developed the behavioral markers. Therefore, it is not surprising that the weighted kappa scores were good for the behavioral markers in Zhang's study. Weighted kappa is statistically similar to two-way mixed, single-measures, consistency ICC, but ICC allows for more than two raters. A third rater in Zhang's study used the original TPOT instrument. Weighted kappa scores between the third rater using the TPOT and the first two raters who used the TBMs were very poor. In the current study, all raters used the TPOT and behavioral markers instrument in order to determine whether one instrument was more sensitive than the other and to determine whether faculty ratings would agree with the ratings provided by the Master Trainers. Faculty in this study chose to complete the behavioral markers instrument prior to the TPOT in every case. Therefore, all raters in the current study essentially used the behavioral markers to inform their ratings for the TPOT. Future studies need to use fully crossed designs with half of the raters completing the behavioral marker instrument and the other half of the raters completing the TPOT without exposure to the behavioral markers. It would also be helpful to use weighted

kappa or two-way mixed, single-measures, consistency ICC so that results may be compared with other studies such as Zhang's.

PRIMARY ANALYSIS

The primary analysis addresses each of the research hypotheses and questions. Key findings are presented here.

INTERRATER AGREEMENT. The first hypothesis proposes that there will be higher interrater agreement on the behavioral markers instrument than the TPOT. However, the ICC values presented do not support Aim 1, Hypothesis 1. Interrater agreement was only greater on the behavioral marker instrument in two out of fifteen rater pairings. Several factors may have influenced these results: faculty expertise in TeamSTEPPs® Teamwork Education, faculty proficiency in evaluating simulations, the level of difficulty in evaluating each team's performance, the order in which the evaluation instruments were completed, and the reliability of the newly developed behavioral marker instrument (Hallgren, 2012).

SENSITIVITY OF OBSERVATIONAL INSTRUMENTS. The second hypothesis proposed that there will be a significant difference between TPOT and behavioral marker assessments indicating greater sensitivity of the behavioral marker (TBM) instrument. A one-way repeated measures ANOVA was calculated comparing results on the TPOT and behavioral markers for all eighteen teams. Since differences between the TPOT and behavioral markers were the within subjects comparison, z scores were used to analyze differences. No significant effect was found ($F(1,17) = .000$, $p = 1.000$) between the TPOT and behavioral markers indicating there is no difference in the scores. Since all raters completed the behavioral markers prior to the TPOT, ratings on the TPOT were likely more accurate due to rater exposure to the behavioral markers. Similar to Zhang's study (2015), the three TeamSTEPPs® Master Trainers in the current study

provided input into refining the behavioral markers. Providing input into the refinement of the behavioral markers may have contributed to the Master Trainers rating agreement on both observational instruments. All other raters were trained on the use of both instruments in order to analyze the amount of agreement between faculty raters who are not TeamSTEPPs® Master Trainers and the Master Trainers. Future study designs need to prevent cross contamination of the rater data by assigning either the TPOT or behavioral markers to each rater group, but not both.

The third hypothesis proposed that correlations between the TPOT, behavioral markers, and teamwork attitudes and perceptions will be higher for the behavioral markers than the TPOT demonstrating greater sensitivity of the behavioral markers. The teamwork perceptions *Situation Monitoring* subscale demonstrated moderate to good negative correlations with the behavioral markers (r = -.520), TPOT (r = -.516), and TPOT subscales of *Communication* (r = -.498), *Leadership* (r = -.551), *and Situation Monitoring* (r = -.520). Although correlations were slightly higher with the behavioral markers instrument, the difference is small. It is likely that the TPOT ratings were contaminated by raters completing the behavioral markers instrument prior to completing the TPOT.

CONVERGENT VALIDITY. The first research question for Aim 2 examined the relationship between the TPOT, behavioral markers, teamwork perceptions, attitudes, and the NLN Simulation Checklist in order to assess for convergent validity. The AHRQ used Kirkpatrick's model of training evaluation which states that reactions, learning, behavior, and results of teamwork education can be measured (AHRQ, 2014e). The TPOT, teamwork perceptions, and attitudes are associated with the evaluation of learning. According to Kirkpatrick's model, the TPOT is associated with behaviors, and the NLN Checklist is

associated with evaluating results or outcomes of the simulation. So, theoretically, data from these evaluations should all improve with teamwork education. The Teamwork Nomological Network theorizes there is a relationship between teamwork attitudes and perceptions, observed teamwork behaviors, and overall performance outcomes of teams. In the Teamwork Nomological Network, moderate to strong relationships are expected between teamwork attitudes, perceptions, observed behaviors, and performance measures. As predicted in the Teamwork Nomological Network, a strong positive relationship of large magnitude was found between the TPOT and behavioral markers (r^2 = .87). In addition, the NLN Checklist also demonstrated a strong positive correlation with the TPOT (r^2 = .62) and the behavioral markers (r^2 = .63). Similarly, the relationship between teamwork perceptions and attitudes also demonstrated a strong positive relationship (r^2 = .47) with one another. No psychometric evidence demonstrating the relationship between teamwork attitudes and perceptions was discovered in the literature prior to this study. The observational performance evaluations (TPOT, behavioral markers, NLN Checklist) do not strongly correlate with the teamwork perceptions and attitudes overall scores. However, correlations between all variables and the T-TPQ *Situation Monitoring* subscale were moderate to good (see Table 8). Therefore, convergent validity was in part established between the TPOT, behavioral markers, NLN Checklist, and the teamwork perceptions *Situation Monitoring* subscale, but not with teamwork attitudes or teamwork perceptions overall scores.

While the TPOT, behavioral markers, and the NLN Simulation Checklist were all moderate to strongly correlated, correlations of all variables with teamwork perceptions and attitudes overall scores were low and frequently negative. This finding is not consistent with the TeamSTEPPs® Instructional Framework or the Teamwork Nomological Network and suggests

that there is incongruence between students' perceptions and attitudes regarding teamwork and faculty's observations of teamwork performance. Further study with fully crossed designs to correct for interrater agreement concerns in this study are needed to explore the relationship between observed teamwork performance, teamwork perceptions, and attitudes.

Although baseline correlations between the TPOT, teamwork perceptions, and attitudes were not found in the literature, several studies used teamwork attitudes and perceptions surveys to evaluate the effectiveness of teamwork education programs. Vertino (2014) found improved teamwork attitudes as measured by the T-TAQ post teamwork education. Conversely, Riggall and Smith (2015) saw a significant decrease in the T-TPQ *Leadership* subscale score from pretest to post-test; however, the post-test was conducted six months after the simulation scenario which may have impacted the findings. Beebe et al. (2012) completed the TPOT during eight Rapid Response Team resuscitations and participants completed the teamwork attitudes survey; however, correlations between the TPOT and teamwork attitudes were not done. Clark used items from the teamwork perceptions and attitudes surveys to evaluate an interprofessional teamwork course. Although improvements were observed from pretest to the course completion post-test, correlations were not conducted between teamwork perceptions and attitudes. In the current study, teamwork perceptions and attitudes surveys were only administered after the simulation, therefore it is not clear whether there were improvements after the simulation experience.

In light of interrater agreement concerns in this study, further investigation regarding the convergence of scores between teamwork perceptions, attitudes, and observed teamwork behavior is needed. Results from the study do not indicate there is a clear correlation between teamwork perceptions, attitudes, and observed teamwork performance. Additional study with

larger sample sizes and fully crossed designs for interrater agreement are needed. Data regarding the relationship between attitudes, perceptions, and behaviors will provide needed information regarding convergent validity and facilitate testing future behavioral markers.

DIVERGENT VALIDITY. The second research question for Aim 2 examines the relationship between the TPOT, behavioral markers, and the Clinical Skills Self-Efficacy Scale in order to assess for divergent validity. The Teamwork Nomological Network theorizes that the aforementioned teamwork measures are weakly related to individual characteristics, such as self-efficacy with clinical skills. Results from this sample indicate that self-efficacy with clinical skills (CSES) is strongly and positively correlated with the TPOT (r^2 = .32) and behavioral markers (r^2 = .42). Strong positive correlations between skill self-efficacy (CSES), TPOT, and behavioral markers indicate that self-efficacy with clinical skills is associated with observed teamwork performance; however, the nature of the relationship or cause cannot be determined from this study. Furthermore, as mentioned with other findings, caution must be used in the extrapolation of these findings due to interrater agreement issues. In additional findings, weak correlations were observed between the skill self-efficacy and teamwork perceptions (r^2 = .016) as well as teamwork attitudes (r^2 = .063) which indicates there is little to no relationship between self-efficacy with clinical skills and teamwork perceptions and attitudes. This low correlation is consistent with the theoretical relationship depicted in the Teamwork Nomological Network between teamwork perceptions, attitudes, and skill self-efficacy. The Clinical Skill Self-Efficacy Scale, teamwork perceptions, and attitudes are all survey instruments completed by the study subjects; therefore, they are not subject to the same interrater agreement concerns as the observational instruments.

MTMM CORRELATION MATRIX. The current study does not fully support the theoretical assertions proposed in the MTMM Correlation Matrix. However, several relationships were identified. The TPOT, behavioral markers, NLN Checklist, and the teamwork perceptions *Situation Monitoring* subscale were strongly correlated. Teamwork perceptions and attitudes were correlated with one another and demonstrated little to no relationship to skill self-efficacy which was consistent with the Teamwork Nomological Network. However, instead of demonstrating divergent validity with the TPOT, behavioral markers, and the NLN Checklist, skill self-efficacy was strongly correlated to each observational instrument. Further study is needed using designs with highly reliable interrater agreement in order to assess these relationships with confidence.

ADDITIONAL FINDINGS

The hypotheses and research questions in this study do not address subject and faculty rater satisfaction. Anecdotally, faculty commented that the behavioral marker instrument was much easier to use during the simulation and provided valuable feedback during the debriefing session. Subjects also verbalized learning a great deal throughout the experience. One subject stated

> "I will definitely take a lot away from this experience. I have learned from my mistakes during this simulation and won't forget them once I start working. Great simulation experience! We should do more like these and in small teams like these. It seems more realistic instead of having 4-5 nurses in the room with you. Great job!"

Another subject commented

> "I liked this simulation because it did allow us to get a feel for appropriate and timely communication in an actual hospital setting; we may not have succeeded,

entirely, but we definitely have a better idea of how we should go about communicating with our colleagues and our clients' family members."

Although the findings in this study indicate observational measures of teamwork need to be refined, the subjects asserted the experience was valuable to their learning process and future professional career. Qualitative exploration of the teamwork educational process may provide greater insight regarding how nursing students learn teamwork.

FEASIBILITY

This study presented several challenges. First, behavioral markers must be developed, tested, and refined in a simulation environment prior to implementing them in a research study. Teamwork behavioral markers are specific to the simulation scenario; therefore, a validation process must be repeated for every simulation scenario in which teamwork behaviors will be evaluated. Second, fully crossed designs which require more than one rater to view each simulation is not always feasible because it increases the number of faculty/staff or the time commitment of faculty/staff needed. In order to complete 18 simulations, five faculty each conducted three to four simulations with the assistance of a Simulation Specialist staff for a time commitment of about ten hours. Rating additional groups would increase the time commitment per faculty in addition to the time it takes for rater training. Multiple raters need to rate the same teams in order to conduct ICCs for interrater agreement. Lastly, students were asked to complete three surveys in addition to the brief demographic form. These surveys were completed after the simulation but prior to the debriefing which delayed the debriefing by approximately 30 to 45 minutes in some cases.

LIMITATIONS

The study utilized a small sample of 18 teams which consisted of a total of 54 subjects at one mid-sized University in Southeast Texas; therefore, findings are limited to this geographical and suburban area, and further limited due to small sample size. Additionally, the method in which Interrater Agreement was derived limits the generalizability of the study findings. Single measures, randomized, absolute agreement Intraclass Correllation Coefficients (ICC) were calculated to analyze interrater agreement. The study design was not fully crossed, meaning that raters were not grouped to always rate the same teams. Rater pairs were analyzed for interrater agreement. Therefore, the design methodology does not meet all the criteria for measuring single measures or average measures ICC which consequently decreases the reliability of the interrater estimates. Additional variables such as the level of simulation, teaching, and TeamSTEPPs® experience of each rater need to be controlled for as much as possible. Research questions and hypotheses based on the observational scores provided by raters need to be interpreted with caution.

FUTURE RESEARCH RECOMMENDATIONS

The development of a Multimethod-Multitrait Correlation Matrix will facilitate the testing of newly developed behavioral markers. Convergent validity in this study was only supported in part by moderate to good positive correlations between the TPOT, behavioral markers, and NLN Checklist but not with teamwork perceptions and attitudes overall scores. Correlations between teamwork observational scores and the NLN Checklist with the teamwork perceptions Situation Monitoring subscale were moderate but negative. However, teamwork perceptions and attitudes overall scores and subscale scores demonstrated moderate to high correlations with one another. The relationship between teamwork attitudes, perceptions, and

observed teamwork behaviors needs to be explored further especially since perceptions and attitudes are often used to measure the learning after teamwork education. However, raters using the TPOT should not be exposed to the behavioral marker instrument in order to prevent contamination of the ratings. Divergent validity not was not established since skill self-efficacy was strongly correlated with TPOT and behavioral markers. However, skill self-efficacy was weakly correlated with teamwork perceptions and attitudes. Further investigation of this relationship using fully crossed designs to detect interrater agreement issues is needed. Finally, studies dependent upon rater agreement need to utilize a completely crossed design in order to accurately calculate ICC. It would also be helpful to separate TeamSTEPPs® Master Trainers from faculty raters in order to detect rating differences between Master Trainers and faculty receiving training. Furthermore, it would be helpful for researchers to use a consistent method for analyzing interrater agreement so that results between studies can be easily compared. Qualitative research methods may also provide valuable insight regarding how nursing students learn and apply teamwork concepts to simulated and actual patient care situations.

SUMMARY

The study examined the relationships between new and established instruments used to determine the outcomes of teamwork education. Several methods are commonly used to evaluate the effectiveness of teamwork education programs: observed teamwork performance scales, teamwork attitudes, and teamwork perceptions. Theoretically, each of these measures should be able to detect changes in teamwork. However, correlations between observed teamwork performance, teamwork attitudes, and teamwork perceptions overall scores were weak in this study. No study was identified with which results could be compared. Several factors may have affected these findings: rater training and experience with simulation, amount of

faculty teaching experience, ambiguity of the behavioral marker instrument, contamination of the TPOT results due to exposure to the behavioral markers, and varying student performance creating difficulty with evaluation. Conversely, self-efficacy with clinical skills demonstrated low correlations with teamwork attitudes and perceptions as predicted, but high correlations with observed teamwork scores. High correlations between clinical skill self-efficacy and observed teamwork performance scores indicate that clinical skill self-efficacy may influence teamwork performance in some manner. Additional study of these relationships using fully crossed designs for interrater agreement are needed in order to clarify the MTMM Correlation Matrix and provide a framework to assess future teamwork behavioral markers.

Appendix A

utmb Health

Working together to work wonders.®

Institutional Review Board
301 University Blvd.
Galveston, TX 77550-0158
409.266.9475

09-Sep-2015

MEMORANDUM

TO: Leann Chisholm
Grad School Biomedical Science GSBS9999

FROM:
Institutional Review Board, Chairman

RE: Exempt from IRB Review

IRB #: IRB # 15-0210

TITLE: Validating Targeted Behavioral Markers for Teamwork Skill and Performance Outcomes in Simulation

The UTMB Institutional Review Board (IRB) reviewed the above-referenced research project and determined this request met the criteria for exemption from review by the IRB in accordance with the 45 CFR 46.101(b). This determination was made on **09-Sep-2015**.

Further review of this project by the IRB is not required unless the protocol changes in the use of human subjects. In that case, the project must be resubmitted to the IRB for review. Please inform the IRB when this research project is completed.

If you have any questions, please do not hesitate to contact the IRB office at 409-266-9475.

Exemption Category

1 and 2
Lamar University is recognized as the recruitment site engaged in research procedures.

Exempt Categories

Category 1: Research conducted in established or commonly accepted educational settings, involving normal educational practices, such as:

i. research on regular and special education instructional strategies, or
ii. research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.

Category 2: Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:

i. information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and
ii. any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

Category 3: Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior that is not exempt under paragraph (b)(2) of this section, if:

i. the human subjects are elected or appointed public officials or candidates for public office; or
ii. Federal statute(s) require(s) without exception that the confidentiality of the personally identifiable information will be maintained throughout the research and thereafter.

Category 4: Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

Category 5: Research and demonstration projects which are conducted by or subject to the approval of Department or Agency heads, and which are designed to study, evaluate, or otherwise examine:

i. public benefit or service programs;
ii. procedures for obtaining benefits or services under those programs;
iii. possible changes in or alternatives to those programs or procedures; or
iv. possible changes in methods or levels of payment for benefits or services under those programs.

Category 6: Taste and food quality evaluation and consumer acceptance studies if:

i. wholesome foods without additives are consumed; or
ii. a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural chemical or environmental contaminant at or below the level found to be safe, by the Food and Drug Administration or approved by the Environmental Protection Agency or the Food Safety and Inspection Service of the U.S. Department of Agriculture.

Appendix B
RESEARCH CONSENT FORM

You are being asked to participate as a subject in the research project entitled, **Validating Targeted Behavioral Markers for Teamwork Skill and Performance Outcomes in Simulation**, under the direction of LeAnn Chisholm MSN, RN, CEN.

PURPOSE OF THE STUDY

The purpose of this study is to test the reliability of behaviors identified to assist faculty in the evaluation of nursing students' teamwork performance. Nurses are expected to enter the profession with the ability to function as a member of the healthcare team. The Institute of Medicine and Quality and Safety Education for Nurses asserts that teamwork is an essential competency for healthcare providers because a lack of teamwork leads to errors in patient care. Therefore, the Preceptorship course has added a teamwork simulation and one hour of team training to the Clinical Evaluation Tool for this course. Students' teamwork behaviors will be evaluated during the simulation scenario.

You are being asked to participate because you are a senior level baccalaureate nursing student enrolled in the final clinical capstone (NURS 4381: Preceptorship) course.

PROCEDURES RELATED ONLY TO THE RESEARCH

Faculty will evaluate teamwork performance of groups of three nursing students during a simulation scenario. These teamwork behaviors were developed in preliminary work and will be tested to determine their reliability and validity in two simulation scenarios.

PROCEDURES NOT RELATED TO THIS RESEARCH (i.e., standard of care)

Students will be randomly assigned to teams consisting of three nursing students. Faculty will assign you to a simulation date which fits your schedule. All students will participate in a one hour team training session. Before the simulation day, each student will be asked to prepare to care for a simulated patient with pancreatitis. Students will be oriented to the room, mannequin/simulator, equipment, and provided with a patient report. Each team will have up to 30 minutes to complete the scenario. Faculty will evaluate team behaviors and completion of specific interventions on a checklist. After the scenario, students will be asked to complete a demographic form, two surveys related to teamwork attitudes and perceptions, and a survey involving self confidence with clinical skills. A recording of the scenario will be used to review teamwork behaviors exhibited and areas for improvement. The total time commitment of each student will be approximately two and one half hours.

RISKS OF PARTICIPATION

The potential risks from participation in the study are minimal since participation in simulation is part of the usual nursing educational process; however, there is a possible risk from loss of confidentiality that may arise from participation in the project. Your name will not be associated with your individual or team evaluation scores and survey data. Each subject and team will receive a code number which will be maintained in a locked cabinet in the Principle Investigator's office. This will be the only link between you and your data.

NUMBER OF SUBJECTS PARTICIPATING AND THE DURATION OF YOUR PARTICIPATION

The anticipated number of subjects involved in the study will be 60 senior baccalaureate nursing students which will be divided into approximately 20 teams. The length of time for your participation is four hours.

BENEFITS TO THE SUBJECT

The direct benefits to you may include improved teamwork and confidence in nursing care abilities.

BENEFITS TO SOCIETY

When students learn how to effectively communicate and work in a team environment, patient outcomes improve and patients benefit. Providing teamwork education along with high fidelity simulation practice will provide advanced training to novice practitioners which may also benefit employers.

SAFE WITHDRAWAL FROM THE STUDY

You have the option to withdraw from the study at any time without penalty. Since the simulation scenarios and team training are part of the clinical capstone (Preceptorship) course, all students must complete this course requirement. Participation involves completing demographic and survey instruments and allowing data to be used in analysis of this research. Students who choose not to participate will be grouped together to participate in simulations.

REIMBURSEMENT FOR EXPENSES

There will be no reimbursement for participation in this study.

COMPENSATION FOR RESEARCH RELATED INJURY

You will be responsible for paying any costs related to illnesses and medical events not associated with this study. There are no plans to provide other forms of compensation. However, you are not waiving any of your legal rights by participating in this study. Questions about compensation may be
directed to the Principle Investigator.

COSTS OF PARTICIPATION

There are no monetary costs to you for participating in this research study.

PROCEDURES FOR WITHDRAWAL

If you decide to withdraw from the study, simply notify your faculty member and you will not be required to complete the surveys. Additionally, your data will not be used in the data analysis.

ADDITIONAL INFORMATION

1. If you have any questions, concerns or complaints before, during or after the research study, or if you need to report a research related injury or adverse reaction (bad side effect), you should immediately contact LeAnn Chisholm at 409-880-8862 or via email at leann.chisholm@lamar.edu.

2. Your participation in this study is completely voluntary and you have been told that you may refuse to participate or stop your participation in this project at any time without penalty or loss of benefits and without jeopardizing your medical care at UTMB. If you decide to stop your participation in this project and revoke your authorization for the use and disclosure of your health information, UTMB may continue to use and disclose your information in some instances. This would include any health information that was used or disclosed prior to your decision to stop participation and needed in order to maintain the integrity of the research study. If there are significant new findings or we get any information that might change your mind about participating, we will give you the information and allow you to reconsider whether or not to continue.

3. If you have any complaints, concerns, input or questions regarding your rights as a subject participating in this research study or you would like more information, you may contact the Institutional Review Board Office, at (409) 266-9475.

The purpose of this research study, procedures to be followed, risks and benefits have been explained to you. You have been allowed to ask questions and your questions have been answered to your satisfaction. You have been told who to contact if you have additional questions. You have read this consent form and voluntarily agree to participate as a subject in this study. You are free to withdraw your consent, including your authorization for the use and disclosure of your health information, at any time. You may withdraw your consent by notifying LeAnn Chisholm at 409-880-8862. You will be given a copy of the consent form you have signed.

Informed consent is required of all persons in this project. Whether or not you provide a signed informed consent for this research study will have no effect on your current or future relationship with UTMB.

Signature of Subject	Date
Signature of Authorized Representative	Date

Description of Authorized Representative's Authority to Act for Subject

Using language that is understandable and appropriate, I have discussed this project and the items listed above with the subject.

Date	Signature of Person Obtaining Consent

Appendix C

Outline of TeamSTEPPs® Education for Nursing Students

Teaching Topics	Strategies	Illustration Method
TeamSTEPPs® overview		Video (2 min. 22 sec.)
Reducing the risk of error		Sue Sheridan (9 min 49 sec.)
Team structure		Discussion
Communication	SBAR	SBAR video (1 min 35 sec.)
	Call-out	Call-out video (18 sec.)
	"I pass the baton"	Handoff video (1min 14 sec.)
Leadership	Brief	Brief video (36 sec.)
	Huddle	Huddle video (31 sec.)
	Debrief	Debrief video (24 sec.)
Situation monitoring	STEP	STEP video (38 sec.)
	Cross-monitoring	Cross-monitoring video (18 sec.)
	I'm SAFE Checklist	Discussion and checklist
Mutual support	Feedback	Feedback video (29 sec.)
	Task Assistance	Discussion
	Advocacy and Assertion	Discussion
	Two-challenge rule	Discussion
	CUS	CUS video (10 sec.)
TeamSTEPPs® outcomes		Video (5 min. 18 sec.)

Appendix D

TeamSTEPPS® 2.0

Team Performance Observation Tool

Date: _____
Unit/Department: _____
Team: _____
Shift: _____

Rating Scale *Please comment if 1 or 2.*
1 = Very Poor
2 = Poor
3 = Acceptable
4 = Good
5 = Excellent

1. Team Structure	**Rating**
a. Assembles a team	
b. Assigns or identifies team members' roles and responsibilities	
c. Holds team members accountable	
d. Includes patients and families as part of the team	
Comments: **Overall Rating – Team Structure**	
2. Communication	**Rating**
a. Provides brief, clear, specific, and timely information to team members	
b. Seeks information from all available sources	
c. Uses check-backs to verify information that is communicated	
d. Uses SBAR, call-outs, and handoff techniques to communicate effectively with team members	
Comments: **Overall Rating – Communication**	
3. Leadership	**Rating**
a. Identifies team goals and vision	
b. Uses resources efficiently to maximize team performance	
c. Balances workload within the team	
d. Delegates tasks or assignments, as appropriate	
e. Conducts briefs, huddles, and debriefs	
f. Role models teamwork behaviors	
Comments: **Overall Rating – Leadership**	
4. Situation Monitoring	**Rating**
a. Monitors the status of the patient	
b. Monitors fellow team members to ensure safety and prevent errors	
c. Monitors the environment for safety and availability of resources (e.g., equipment)	
d. Monitors progress toward the goal and identifies changes that could alter the plan of care	
e. Fosters communication to ensure that team members have a shared mental model	
Comments: **Overall Rating – Situation Monitoring**	
5. Mutual Support	**Rating**
a. Provides task-related support and assistance	
b. Provides timely and constructive feedback to team members	
c. Effectively advocates for patient safety using the Assertive Statement, Two-Challenge Rule, or CUS	
d. Uses the Two-Challenge Rule or DESC Script to resolve conflict	
Comments: **Overall Rating – Mutual Support**	
TEAM PERFORMANCE RATING	

Appendix E

Targeted Behavioral Markers

Simulation: Pancreatitis Hemodynamic and Respiratory Instability

Triggering Events	**Associated TPOT Items**	**Targeted Behavioral Markers**
Nurses enter patient room and wash hands. Introduce members of team	1.a. Assembles a team.	( ) All wash hands or one nurse reminds the other to wash hands if not done ( ) Members introduce self and/or one another to patient and patient's family (or remind one another). ***Check this if done without reminding.**
Nurses enter room, review patient information, and develop a plan.	1.b. Assigns team members' roles and responsibilities. 3.a. Identifies team goals and vision. 3.b. Uses resources efficiently to maximize team performance. 3.c. Balances workload within the team. 3.d. Delegates tasks or assignments, as appropriate. 3.e. Conducts briefs, huddles, and debriefs.	( ) Nurses brief to develop a plan (divide tasks). ( ) Nurses huddle to discuss/modify the plan.
Nurses conduct physical assessment and review chart. Patient confused and stating "I am in pain."	1.c. Holds team members accountable. 2.a. Provides brief, clear, specific, and timely information to team members 2.b. Seeks information from all available sources. 2.c. Uses check-backs to verify information that is communicated. 2.d. Uses SBAR, call-outs, and handoff techniques to communicate effectively with team members. 3.f. Role models teamwork behaviors. 4.a. Monitors the status of the patient.	( ) Nurse gathers physical assessment data and discuss with other nurse. ( ) Second nurses verbalizes receiving information. ( ) Requests additional information if needed.

	1.c. Holds team members accountable. 2.a. Provides brief, clear, specific, and timely information to team members 2.b. Seeks information from all available sources. 2.c. Uses check-backs to verify information that is communicated. 2.d. Uses SBAR, call-outs, and handoff techniques to communicate effectively with team members. 3.f. Role models teamwork behaviors. 4.a. Monitors the status of the patient.	( ) Nurse gathers information from the chart (labs, orders, etc.) and discuss with other nurse
	3.e. Conducts briefs, huddles, and debriefs. 4.a. Monitors the status of the patient. 4.d. Monitors progress toward the goal and identifies changes that could alter the plan of care. 4.e. Fosters communication to ensure that team members have a shared mental model.	( ) Nurses discuss situation (identify breathing difficulty, BP decrease and HR elevation) which needs to be communicated to PCP.
Patient states "I am tired and short of breath."	2.a. Provides brief, clear, specific, and timely information to team members. 2.b. Seeks information from all available sources. 2.d. Uses SBAR, call-outs, and handoff techniques to communicate effectively with team members. 4.d. Monitors progress toward the goal and identifies changes that could alter the plan of care. 4.e. Fosters communication to ensure that team members have a shared mental model.	( ) One nurse communicates breathing difficulty, low BP and high HR to PCP.
Nurse receives orders	2.c. Uses check-backs to verify information that is communicated.	( ) Reads back verbal/phone order.

	1.b. Assigns or identifies team members' roles and responsibilities. 3.b. Uses resources efficiently to maximize team performance. 4.c. Monitors the environment for safety and availability of resources. 5.a. Provides task-related support and assistance.	( ) Nurses coordinate completing orders (increasing oxygen, administering Calcium gluconate, fluid bolus, and increasing IV rate).
	4.b. Monitors fellow team members to ensure safety and prevent errors. 5.a. Provides task related support and assistance. 5.b. Provides timely and constructive feedback to team members. 5.c. Effectively advocates for patient safety using the Assertive Statement, Two-Challenge Rule, or CUS 5.d. Uses the Two-Challenge Rule or DESC Script to resolve conflict.	( ) Nurses discuss safety of administering Calcium gluconate IVP (dose, route, rate)
	4.a. Monitors the status of the patient. 4.d Monitors progress toward goal and identifies changes that could alter the plan of care. 4.e. Fosters communication to ensure that team members have a shared mental model.	( ) Nurse communicates patient evaluation (vital signs) after interventions to other nurse before calling physician.
	1.d. Includes patients and families as part of the team. 4.e. Fosters communication to ensure that team members have a shared mental model.	( ) Nurses discuss plan of care with patient and family
Orders implemented and patient remains short of breath with oxygen saturation 91%	3.b. Uses resources efficiently to maximize team performance. 3.e. Conducts briefs, huddles, and debriefs. 4.a. Monitors the status of the patient.	( ) Nurses discuss patient current status and determine that additional help is needed. ( ) Call for additional support (RT, PCP, Supervisor, Charge Nurse).

	4.d. Monitors progress toward the goal and identifies changes that could alter the plan of care. 4.e. Fosters communication to ensure that team members have a shared mental model.	

Appendix F: TeamSTEPPs® Teamwork Perceptions Questionnaire (T-TPQ)

TeamSTEPPS® 2.0

TeamSTEPPS Teamwork Perceptions Questionnaire (T-TPQ)

Instructions: Please respond to the questions below by placing a check mark (√) in the box that corresponds to your level of agreement from *Strongly Agree* to *Strongly Disagree*. Please select only one response for each question.

		Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Team Structure						
1.	The skills of staff overlap sufficiently so that work can be shared when necessary.					
2.	Staff are held accountable for their actions.					
3.	Staff within my unit share information that enables timely decision making by the direct patient care team.					
4.	My unit makes efficient use of resources (e.g., staff supplies, equipment, information).					
5.	Staff understand their roles and responsibilities.					
6.	My unit has clearly articulated goals.					
7.	My unit operates at a high level of efficiency.					
Leadership						
8.	My supervisor/manager considers staff input when making decisions about patient care.					
9.	My supervisor/manager provides opportunities to discuss the unit's performance after an event.					
10.	My supervisor/manager takes time to meet with staff to develop a plan for patient care.					
11.	My supervisor/manager ensures that adequate resources (e.g., staff, supplies, equipment, information) are available.					
12.	My supervisor/manager resolves conflicts successfully.					
13.	My supervisor/manager models appropriate team behavior.					
14.	My supervisor/manager ensures that staff are aware of any situations or changes that may affect patient care.					

PLEASE CONTINUE TO THE NEXT PAGE

		Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Situation Monitoring						
15.	Staff effectively anticipate each other's needs.					
16.	Staff monitor each other's performance.					
17.	Staff exchange relevant information as it becomes available.					
18.	Staff continuously scan the environment for important information.					
19.	Staff share information regarding potential complications (e.g., patient changes, bed availability).					
20.	Staff meets to reevaluate patient care goals when aspects of the situation have changed.					
21.	Staff correct each other's mistakes to ensure that procedures are followed properly.					
Mutual Support						
22.	Staff assist fellow staff during high workload.					
23.	Staff request assistance from fellow staff when they feel overwhelmed.					
24.	Staff caution each other about potentially dangerous situations.					
25.	Feedback between staff is delivered in a way that promotes positive interactions and future change.					
26.	Staff advocate for patients even when their opinion conflicts with that of a senior member of the unit.					
27.	When staff have a concern about patient safety, they challenge others until they are sure the concern has been heard.					
28.	Staff resolve their conflicts, even when the conflicts have become personal.					

PLEASE CONTINUE TO THE NEXT PAGE

		Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Communication						
29.	Information regarding patient care is explained to patients and their families in lay terms.					
30.	Staff relay relevant information in a timely manner.					
31.	When communicating with patients, staff allow enough time for questions.					
32.	Staff use common terminology when communicating with each other.					
33.	Staff verbally verify information that they receive from one another.					
34.	Staff follow a standardized method of sharing information when handing off patients.					
35.	Staff seek information from all available sources.					

Appendix G: TeamSTEPPs® Teamwork Attitudes Questionnaire (T-TAQ)

TeamSTEPPS™ Teamwork Attitudes Questionnaire

The purpose of this survey is to measure your impressions of various components of teamwork as it relates to patient care and safety.

Instructions: Please respond to the questions below by placing a check mark (√) in the box that corresponds to your level of agreement from *Strongly Disagree* to *Strongly Agree*. Please select only one response for each question.

TeamSTEPPS

		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Team Structure						
1.	It is important to ask patients and their families for feedback regarding patient care.					
2.	Patients are a critical component of the care team.					
3.	This facility's administration influences the success of direct care teams.					
4.	A team's mission is of greater value than the goals of individual team members.					
5.	Effective team members can anticipate the needs of other team members.					
6.	High-performing teams in health care share common characteristics with high-performing teams in other industries.					
Leadership						
7.	It is important for leaders to share information with team members.					
8.	Leaders should create informal opportunities for team members to share information.					
9.	Effective leaders view honest mistakes as meaningful learning opportunities.					
10.	It is a leader's responsibility to model appropriate team behavior.					
11.	It is important for leaders to take time to discuss with their team members plans for each patient.					
12.	Team leaders should ensure that team members help each other out when necessary.					

PLEASE CONTINUE TO THE NEXT PAGE

		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Situation Monitoring						
13.	Individuals can be taught how to scan the environment for important situational cues.					
14.	Monitoring patients provides an important contribution to effective team performance.					
15.	Even individuals who are not part of the direct care team should be encouraged to scan for and report changes in patient status.					
16.	It is important to monitor the emotional and physical status of other team members.					
17.	It is appropriate for one team member to offer assistance to another who may be too tired or stressed to perform a task.					
18.	Team members who monitor their emotional and physical status on the job are more effective.					
Mutual Support						
19.	To be effective, team members should understand the work of their fellow team members.					
20.	Asking for assistance from a team member is a sign that an individual does not know how to do his/her job effectively.					
21.	Providing assistance to team members is a sign that an individual does not have enough work to do.					
22.	Offering to help a fellow team member with his/her individual work tasks is an effective tool for improving team performance.					
23.	It is appropriate to continue to assert a patient safety concern until you are certain that it has been heard.					
24.	Personal conflicts between team members do not affect patient safety.					

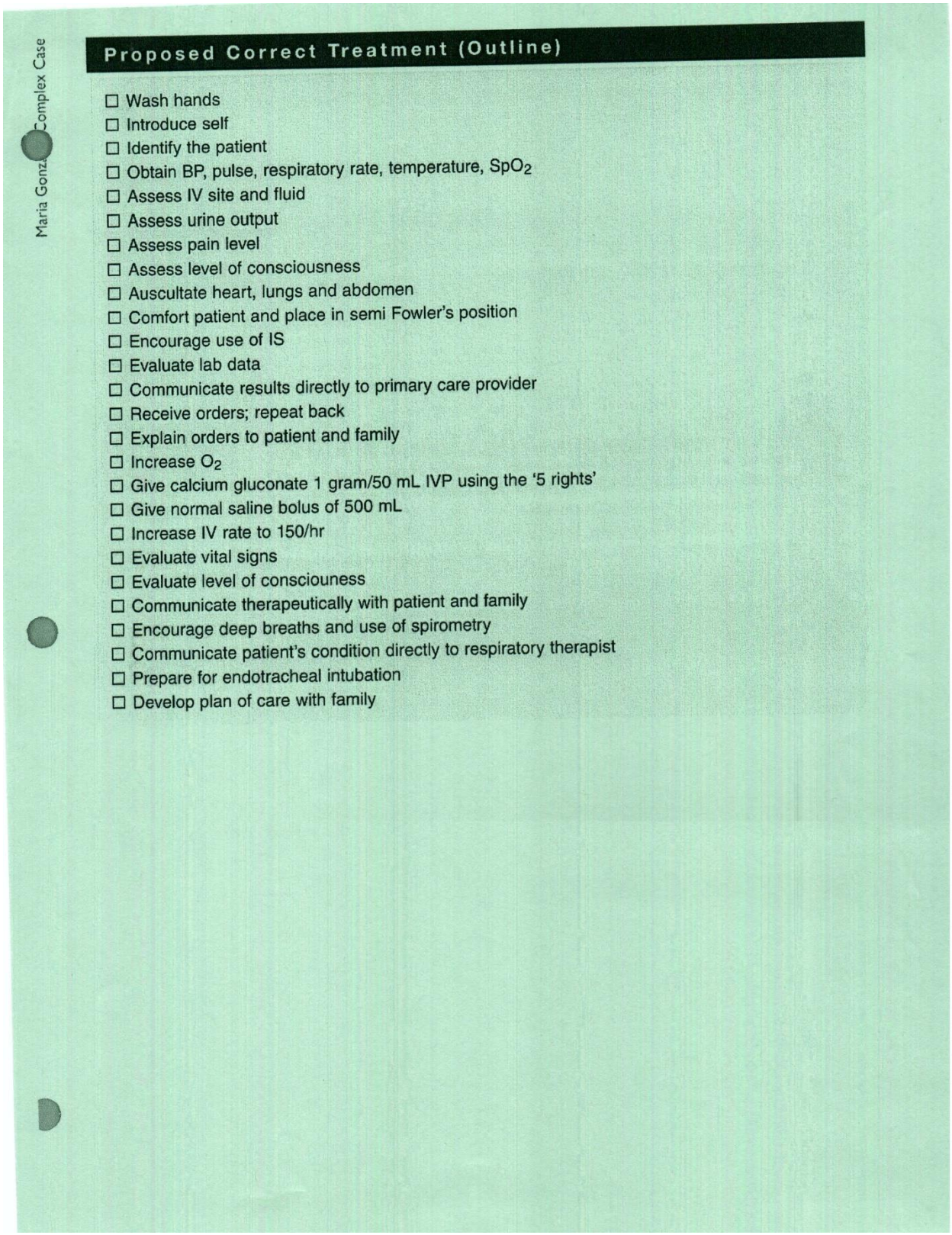
PLEASE CONTINUE TO THE NEXT PAGE

		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Communication						
25.	Teams that do not communicate effectively significantly increase their risk of committing errors.					
26.	Poor communication is the most common cause of reported errors.					
27.	Adverse events may be reduced by maintaining an information exchange with patients and their families.					
28.	I prefer to work with team members who ask questions about information I provide.					
29.	It is important to have a standardized method for sharing information when handing off patients.					
30.	It is nearly impossible to train individuals how to be better communicators.					

Please provide any additional comments in the space below.

Thank you for your participation!

Appendix H: Simulation Checklists Pancreatitis Hemodynamic instability



Maria Gonza Complex Case

Proposed Correct Treatment (Outline)

- ☐ Wash hands
- ☐ Introduce self
- ☐ Identify the patient
- ☐ Obtain BP, pulse, respiratory rate, temperature, SpO_2
- ☐ Assess IV site and fluid
- ☐ Assess urine output
- ☐ Assess pain level
- ☐ Assess level of consciousness
- ☐ Auscultate heart, lungs and abdomen
- ☐ Comfort patient and place in semi Fowler's position
- ☐ Encourage use of IS
- ☐ Evaluate lab data
- ☐ Communicate results directly to primary care provider
- ☐ Receive orders; repeat back
- ☐ Explain orders to patient and family
- ☐ Increase O_2
- ☐ Give calcium gluconate 1 gram/50 mL IVP using the '5 rights'
- ☐ Give normal saline bolus of 500 mL
- ☐ Increase IV rate to 150/hr
- ☐ Evaluate vital signs
- ☐ Evaluate level of consciouness
- ☐ Communicate therapeutically with patient and family
- ☐ Encourage deep breaths and use of spirometry
- ☐ Communicate patient's condition directly to respiratory therapist
- ☐ Prepare for endotracheal intubation
- ☐ Develop plan of care with family

Appendix I: Clinical Skills Self Efficacy Scale (CSES)

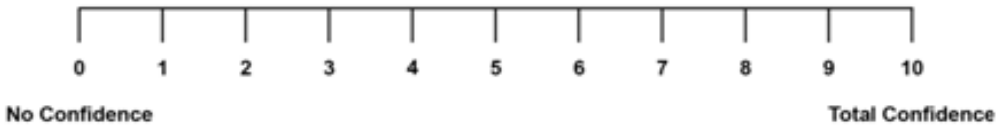
Clinical Skills Self-Efficacy Scale

DIRECTIONS: This questionnaire should take no more than 10-15 minutes to complete.

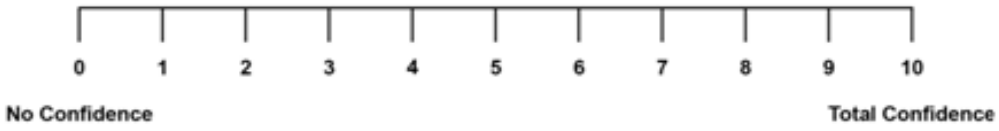
Each of the statements below is written so nursing students can describe their perceptions of their confidence in performing certain skills that they are routinely expected to do in their clinical settings.

Please **circle the number** that identifies how confident you are **right now** of your ability to perform each of the behaviors. Remember there is no right, or wrong answers but it is very important that you answer the questions honestly.

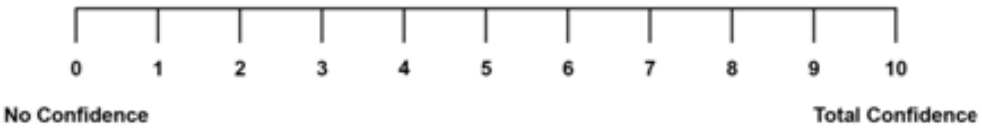
1. How confident are you **right now** that you can independently administer an intramuscular injection?



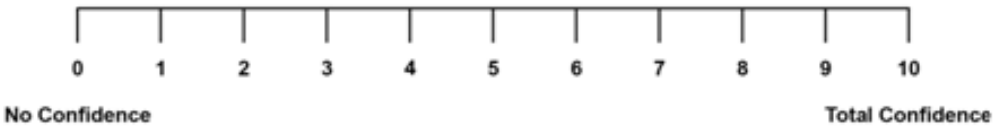
2. How confident are you **right now** that you can independently administer an insulin injection?



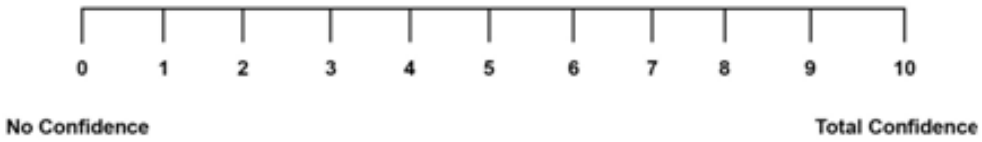
3. How confident are you **right now** that you can independently change a dressing maintaining sterile technique?



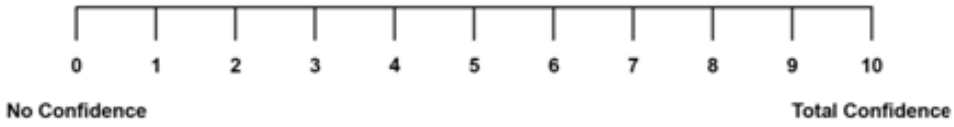
4. How confident are you **right now** that you can independently insert a Foley catheter using sterile technique?



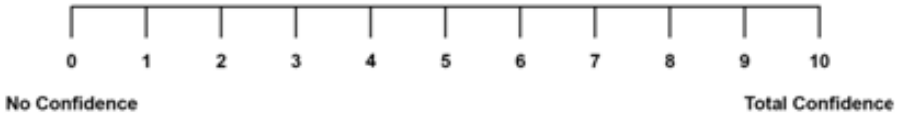
5. How confident are you **right now** that you can insert a nasogastric tube with correct placement?



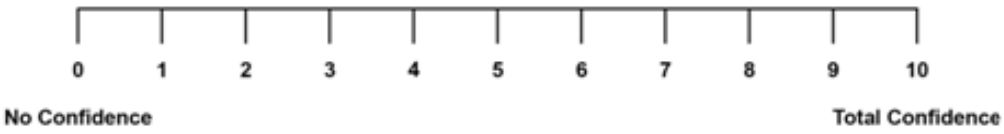
6. How confident are you **right now** that can independently start an intravenous line?



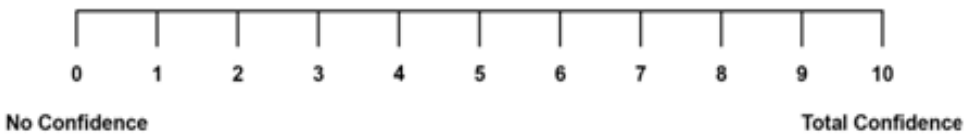
7. How confident are you **right now** that you can correctly transfer an immobile patient from bed to chair using correct technique?



8. How confident are you **right now** that you can independently hang an intravenous piggyback medicine and program the pump accurately?



9. How confident are you **right now** that you can administer a tube feeding through a PEG tube using correct technique?



Before finishing this questionnaire, please fill in all of the blank spaces in this section:

1. What is your age? _____

2. Male _____ Female _____

3. Clinical course currently enrolled in _____

4. Have you ever administered an intramuscular injection?

   Yes _____ No _____

5. Have you ever changed a dressing using sterile technique?

   Yes _____ No _____

6. Have you ever inserted a Foley catheter?

Yes _____ No_____

Thank You for completing this questionnaire!

Today's Date _____

Appendix J

Summary of Observational Instrument Scores

TEAM	TPOT without TBM Overall	TPOT without TBM Mean	TPOT with TBM Overall	TPOT with TBM Percent	NLN Overall	NLN Percent
1	161	2.36	23	45.10	41	52.56
2	245	3.53	44	86.27	55	70.51
3	187	2.72	28	54.90	49	62.82
4	169	2.43	27	52.94	44	56.41
5	236	3.42	40	78.43	51	65.38
6	283	4.10	42	82.35	49	62.82
7	230	3.32	33	64.71	53	67.95
8	231	3.33	31	60.78	48	61.54
10	195	2.81	31	60.78	40	51.28
11	267	3.85	42	82.35	54	69.23
12	260	3.73	41	80.39	62	79.49
13	233	3.37	36	70.59	51	65.38
14	278	4.03	44	86.27	71	91.03
15	242	3.51	41	80.39	53	67.95
16	163	2.36	25	49.02	34	43.59
17	222	3.20	38	74.51	55	70.51
18	275	3.93	44	86.27	64	82.05
19	236	3.41	37	72.55	47	60.26

Appendix K

Summary of Overall and Mean Survey Scores by Team

TEAM	TPQ Overall	TPQ Item Mean	TAQ	TAQ Mean	CSES Overall	CSES Mean
1	500	4.76	378	4.20	177	6.56
2	455	4.33	412	4.58	245	9.07
3	469	4.47	389	4.32	165	6.11
4	445	4.24	401	4.46	201	7.44
5	417	3.97	410	4.56	244	9.04
6	490	4.67	387	4.30	228	8.44
7	440	4.19	426	4.73	210	7.78
8	488	4.65	380	4.22	214	7.93
10	396	3.77	433	4.81	170	6.30
11	438	4.17	397	4.41	246	9.11
12	435	4.14	428	4.76	222	8.22
13	407	3.88	397	3.94	190	7.04
14	419	3.99	381	4.41	215	7.96
15	462	4.40	426	4.73	219	8.11
16	512	4.88	418	4.64	231	8.56
17	508	4.84	420	4.67	231	8.56
18	442	4.21	421	4.68	233	8.63
19	455	4.33	424	4.71	238	8.81

Appendix L

Summary of Overall, Mean, and Percentage Scores for Observational and Survey Instruments by Team

TEAM	TPOT Overall	TPOT Item Mean	TBM Overall	TBM Percent	NLN Overall	NLN Percent	TPQ Overall	TPQ Item Mean	TAQ Overall	TAQ Mean	CSES Overall	CSES Mean
1	161	2.36	23	45.10	41	52.56	500	4.76	378	4.20	177	6.56
2	245	3.53	44	86.27	55	70.51	455	4.33	412	4.58	245	9.07
3	187	2.72	28	54.90	49	62.82	469	4.47	389	4.32	165	6.11
4	169	2.43	27	52.94	44	56.41	445	4.24	401	4.46	201	7.44
5	236	3.42	40	78.43	51	65.38	417	3.97	410	4.56	244	9.04
6	283	4.10	42	82.35	49	62.82	490	4.67	387	4.30	228	8.44
7	230	3.32	33	64.71	53	67.95	440	4.19	426	4.73	210	7.78
8	231	3.33	31	60.78	48	61.54	488	4.65	380	4.22	214	7.93
10	195	2.81	31	60.78	40	51.28	396	3.77	433	4.81	170	6.30
11	267	3.85	42	82.35	54	69.23	438	4.17	397	4.41	246	9.11
12	260	3.73	41	80.39	62	79.49	435	4.14	428	4.76	222	8.22
13	233	3.37	36	70.59	51	65.38	407	3.88	397	3.94	190	7.04
14	278	4.03	44	86.27	71	91.03	419	3.99	381	4.41	215	7.96
15	242	3.51	41	80.39	53	67.95	462	4.40	426	4.73	219	8.11
16	163	2.36	25	49.02	34	43.59	512	4.88	418	4.64	231	8.56
17	222	3.20	38	74.51	55	70.51	508	4.84	420	4.67	231	8.56
18	275	3.93	44	86.27	64	82.05	442	4.21	421	4.68	233	8.63
19	236	3.41	37	72.55	47	60.26	455	4.33	424	4.71	238	8.81

Appendix M

Correlations Between TPOT, Behavioral Markers (TBM), NLN Checklist, T-TPQ Overall and Subscale Scores

	NLN Checklist	TBM	TPOT Overall	TPOT (TS Subscale)	TPOT (C Subscale)	TPOT (L Subscale)	TPOT (SM Subscale)	TPOT (MS Subscale)	T-TPQ Overall	TPQ-TS Subscale	TPQ-L Subscale	TPQ-SM Subscale	TPQ-MS Subscale	TPQ-C Subscale
NLN Checklist	1	.791**	.789**	.541*	.748**	.740**	.842**	.732**	-.357	-.188	-.214	-.490*	-.275	-.436
TBM		1	.933**	.714**	.904**	.933**	.847**	.871**	-.354	-.166	-.221	-.520*	-.243	-.435
TPOT Overall			1	.777**	.946**	.968**	.945**	.957**	-.315	-.151	-.146	-.516*	-.226	-.365
TPOT (TS Subscale)				1	.616**	.709**	.686**	.665**	-.156	-.068	-.035	-.315	-.054	-.214

	NLN Checklist	TBM	TPOT Overall	TPOT (TS Subscale)	TPOT (C Subscale)	TPOT (L Subscale)	TPOT (SM Subscale)	TPOT (MS Subscale)	T-TPQ Overall	TPQ-TS Subscale	TPQ-L Subscale	TPQ-SM Subscale	TPQ-MS Subscale	TPQ-C Subscale
TPOT (C Subscale)					1	.921**	.850**	.941**	-.346	-.203	-.195	-.498*	-.316	-.338
TPOT (L Subscale)						1	.883**	.893**	-.309	-.103	-.088	-.551*	-.231	-.398
TPOT (SM Subscale)							1	.908**	-.342	-.188	-.201	-.520*	-.216	-.406
TPOT (MS Subscale)								1	-.260	-.134	-.150	-.427	-.188	-.260
TPQ Overall									1	.918**	.887**	.905**	.946**	.896**

	NLN Checklist	TBM	TPOT Overall	TPOT (TS Subscale)	TPOT (C Subscale)	TPOT (L Subscale)	TPOT (SM Subscale)	TPOT (MS Subscale)	T-TPQ Overall	TPQ-TS Subscale	TPQ-L Subscale	TPQ-SM Subscale	TPQ-MS Subscale	TPQ-C Subscale
TPQ TS Subscale										1	.841**	.768**	.887**	.695**
TPQ L Subscale											1	.648**	.809**	.763**
TPQ SM Subscale												1	.845**	.838**
TPQ MS Subscale													1	.769**
TPQ C Subscale														1

Appendix N

Correlations Between TPOT, Behavioral Markers (TBM), NLN Checklist, T-TAQ Overall and Subscale Scores

	NLN Checklist	TBM	TPOT Overall	TPOT (TS Subscale)	TPOT (C Subscale)	TPOT (L Subscale)	TPOT (SM Subscale)	TPOT (MS Subscale)	T-TAQ Overall	TAQ-TS Subscale	TAQ-L Subscale	TAQ-SM Subscale	TAQ-MS Subscale	TAQ-C Subscale
NLN Checklist	1	.791**	.789**	.541*	.748**	.740**	.842**	.732**	-.015	-.070	-.325	-.243	.208	.042
TBM		1	.933**	.714**	.904**	.933**	.847**	.871**	.191	-.184	-.340	-.259	-.153	-.038
TPOT Overall			1	.777**	.946**	.968**	.945**	.957**	.028	-.165	-.412	-.274	-.091	-.043
TPOT (TS Subscale)				1	.616**	.709**	.686**	.665**	-.024	-.131	-.440	-.303	.098	-.177

	NLN Checklist	TBM	TPOT Overall	TPOT (TS Subscale)	TPOT (C Subscale)	TPOT (L Subscale)	TPOT (SM Subscale)	TPOT (MS Subscale)	T-TAQ Overall	TAQ-TS Subscale	TAQ-L Subscale	TAQ-SM Subscale	TAQ-MS Subscale	TAQ-C Subscale
TPOT (C Subscale)					1	.921**	.850**	.941**	-.053	-.200	-.338	-.306	-.128	-.037
TPOT (L Subscale)						1	.883**	.893**	.152	-.078	-.298	-.233	-.266	-.007
TPOT (SM Subscale)							1	.908**	.048	-.196	-.455	-.250	-.004	-.051
TPOT (MS Subscale)								1	-.091	-.192	-.434	-.198	.009	.037
TAQ Overall									1	.799**	.627**	.809**	.392	.834**

TAQ C Subscale	TAQ MS Subscale	TAQ SM Subscale	TAQ L Subscale	TAQ TS Subscale	
					NLN Checklist
					TBM
					TPOT Overall
					TPOT (TS Subscale)
					TPOT (C Subscale)
					TPOT (L Subscale)
					TPOT (SM Subscale)
					TPOT (MS Subscale)
					T-TAQ Overall
				1	TAQ-TS Subscale
			1	.770**	TAQ-L Subscale
		1	.779**	.765**	TAQ-SM Subscale
	1	-.092	-.340	-.097	TAQ-MS Subscale
1	.035	.773**	.597**	.738**	TAQ-C Subscale

References

Adamson, K. A., Kardong-Edgren, S., & Willhaus, J. (2013). An updated review of published simulation evaluation instruments. *Clinical Simulation in Nursing, 9*(9), e393-e400. http://dx.doi.org/10.1016/j.ecns.2012.09.004.

Agency for Healthcare Research and Quality (2014a). Team performance observation tool: TeamSTEPPS® instructor manual. Agency for Healthcare Research and Quality, Rockville, MD. Retrieved from http://www.ahrq.gov/professionals/education/curriculumtools/teamstepps/instructor/reference/tmpot.html

Agency for Healthcare Research and Quality (2014b). Teamwork perceptions questionnaire: TeamSTEPPS® instructor manual. Agency for Healthcare Research and Quality, Rockville, MD. Retrieved from http://www.ahrq.gov/professionals/education/curriculum-tools/teamstepps/instructor/reference/teampercept.html

Agency for Healthcare Research and Quality (2014c). Teamwork attitudes questionnaire: TeamSTEPPS® instructor manual. Agency for Healthcare Research and Quality, Rockville, MD. Retrieved from http://www.ahrq.gov/professionals/education/curriculum-tools/teamstepps/instructor/reference/teamattitude.html

Agency for Healthcare Research and Quality (2014d). Measurement: TeamSTEPPs® Instructor Manual. Agengy for Healthcare Research and Quality, Rockville, MD. Retrieved from www.ahrq.gov/.../education/curriculum-tools/teamstepps/instructor/fundamentals/module10/slmeasure.pptx

Agency for Healthcare Research and Quality (2014e). TeamSTEPPs® Instructor Guide. Agency for Healthcare Research and Quality, Rockville, MD. Retrieved from http://www.ahrq.gov/professionals/education/curriculum-tools/teamstepps/instructor/fundamentals/index.html

Agency for Healthcare Research and Quality (2014f). 20 Tips to Help Prevent Medical Errors. December 2014. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/patients-consumers/care-planning/errors/20tips/index.html, retrieved Feb. 6, 2016

Baker, D. P. & Salas, E. (1996). Analyzing team performance: In the eye of the beholder? *Military Psychology, 8*(3), 235-245.

Baker, D., Day, R., & Salas, E. (2006). Teamwork as an essential component of high-reliability organizations. *Health Services Research*, *41*(4P2), 1576-1598 23p.

Bandura, A. (1986). *Social foundations of thought and action: a social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Barnsteiner, J., Disch, J., Johnson, J., McGuinn, K., Chappell, K., & Swartwout, E. (2012). Diffusing QSEN competencies across schools of nursing: the AACN/RWJF faculty development institutes. *Journal of Professional Nursing, 29*(2), 68-74. dx.doi.org/10.1016/j.profnurs.2012.12.003

Barton, A.J., Armstrong, G., Prehelm, G., Gelmon, S.B., & Andrus, L.C. (2009). A national delphi to determine developmental progression of quality and safety competencies in nursing education. *Nursing Outlook, 57*(6), 313-322. doi 10.1016/j.outlook.2009.08.003

Brock, D., Abu-Rish, E., Chiu, Chia-Ru, Hammer, D., Wilson, S., Vorvick, L., Blondon, Kaherine, Schaad, D., Liner, D., & Zierler, B. (2013). Interprofessional education in team communication: Working together to improve patient safety. *BMJ Quality and Safety, 22,* 414-423.

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

Billings, D. M. & Halstead, J. A. (2012). *Teaching in nursing: A guide for faculty* (4th ed.). St. Louis, Missouri: Elsevier Saunders.

Bowers, C. A., Baker, D. P., & Salas, E. (1994). The importance of teamrk in the cockpit: The utility of job/task analysis indices for training design. *Military Psychology, 6,* 205-214.

Capella, J., Smith, S., Philp, A., Putnam, T., Gilbert, C., Fry, W., Harvey, E., Wright, A., Henderson, K., Baker, D., Ranson, S., & ReMine, S. (2010). Teamwork training improves the clinical care of trauma patients. *Journal of Surgical Education, 67*(6), 439-443.

Cicchetti, D. V. (1994). Gidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284-290.

Clapper, T. C., & Kong, M. (2012). TeamSTEPPs®: The patient safety tool that needs to be implemented. *Clinical Simulation in Nursing, 8*(8), e367-e373. Doi:10.1016/j.ecns.2011.03.002.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychologicl Bulletin, 52*(4), 281-302.

Cronenwett, L., Sherwood, G., Barnsteiner, J. Disch, J., Johnson, J., Mitchell, P., Sullivan, D. T., & Warren, J. (2007). Quality and safety education for nurses. *Nursing Outlook, 55,* 122-131. Doi:10.1016/j.outlook.2007.02.006

Curl, E. D., Smith, S., Chisholm, L. A., McGee, L. A., & Das, K. (2016). Effectiveness of Integrated Simulation and Clinical Experiences Compared to Traditional Clinical Experiences for Nursing Students. *Nursing Education Perspectives*, *37*(2), 72-77 6p. doi:10.5480/15-1647

Dwyer, D. J., Oser, R. L., Salas, E., & Fowlkes, J. E. (1999). Performance measurement in distributed environments: Initial results and implications for training. *Military Psychology*, *11*(2), 189-215.

Eppich, W., & Cheng, A. (2015). Promoting Excellence and Reflective Learning in Simulation (PEARLS): development and rationale for a blended approach to health care simulation debriefing. *Simulation In Healthcare: Journal Of The Society For Simulation In Healthcare*, *10*(2), 106-115. doi:10.1097/SIH.0000000000000072

Fowlkes, J., Dwyer, D. J., Oser, R. L., & Salas, E. (1998). Event-based approach to training (EBAT). *The International Journal of Aviation Psychology, 8*(3), 209-221.

Graham, M., Milanowski, A., & Miller, J. (2012). Measuring and promoting inter-rater agreement of teacher and principal performace ratings. Online Submission, Center for Educator Compensation Reform.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutor Quant Methods Psychol., 8*(1), 23-34.

Hayden, J. K., Smiley, R. A., Alexander, M., Kardong-Edgren, S., & Jeffries, P. R. (2014). The NCSBN national simulation study: A longitudinal, randomized, controlled study replacing clinical hours with simulation in prelicensure nursingeducation. *Journal of Nursing Regulation*, *5*(2), July 2014 Supplement.

Hirokawa, R.Y., Daub, K., Lovell, E., Smith, S., Davis, A., & Beck, C. (2012). Using a human patient simulator to study the relationship between communication and nursing students' team performance. *Journal of Nursing Education, 51*(11), 647-651.

Institute of Medicine (2001). *Crossing the quality chasm: A new health system for the 21st century*.National Academies Press. Retrieved from http://www.nap.edu/catalog/10027.html

Institute of Medicine (2003). *Health professions education: A bridge to quality.* National Academies Press. Retrieved from http://www.nap.edu/catalog/10681.html

International Civil Aviation Organization (2014). *Safety Report.* Montréal, Canada: International Civil Aviation Organization. Retrieved from http://www.icao.int/safety/Documents/ICAO_2014%20Safety%20Report_final_02042014_web.pdf

James, J. T. (2013). A new, evidence-based estimate of patient harms associated with hospital care. *Journal of Patient Safety, 9*(3), 122-128.

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

King HB, Battles J, Baker DP, et al. (2008). TeamSTEPPS™: Team Strategies and Tools to Enhance Performance and Patient Safety. In: Henriksen K, Battles JB, Keyes MA, et al., editors. Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 3: Performance and Tools). Rockville (MD): Agency for Healthcare Research and Quality (US); 2008 Aug. Available from: http://www.ncbi.nlm.nih.gov/books/NBK43686/

Liaw, S. Y., Scherpbier, A., Rethans, J. & Klainin-Yobas, P. (2012). Assessment for simulation learning outcomes: a comparison of knowledge and self-reported confidence with observed clinical performance. *Nurse Education Today, 32,* e35-e39. Doi 10.1016/jnedt2011.10.006

Malouf, N. & West, S. (2011). Fitting in: A pervasive new graduate nursing need. *Nurse Education Today, 31,* 488-493.

MacDonald, M.B., Bally, J.M., Ferguson, L.M., Murray, B.L., Fowler-Kerry, S.E., & Anonson, J.M.S. (2010). Knowledge of the professional role of others: A key interprofessional competency. *Nurse Education in Practice, 10,* 238-242. doi 10.1016/j.nepr.2009.11.012

Maguire, M. B., Bremmer, M. N., & Yanosky, D. J. (2014). Reliability and validity testing of pilot data from the TeamSTEPPs® performance observation tool. *Journal of Nursing Care, 3*(5), 1-4. http://dx.doi.org/10.4172/2167-1168.1000202.

National League for Nursing (2015). A vision for teaching with simulation. NLN Vision Series: Transforming Nursing Education Leading the Call to Reform. Retrieved from http://www.nln.org/docs/default-source/about/nln-vision-series-%28position- statements%29/vision-statement-a-vision-for-teaching-with-simulation.pdf?sfvrsn=2

Oetker-Black, S. L. & Kreye, J. (2015). Global psychometric evaluation of the clinical skills self-efficacy scale in Moshi, Tanzania, *Nursing Education Perspectives,36*(3), 163-166. Doi 10.5480/13-1256.

Patterson, M. D., Geis, G. L., LeMaster, T., & Wears, R.L. (2013). Impact of a multidisciplinary simulation-based training on patient safety in a paediatric emergency department. *BMJ Quality & Safety*, 22, 383-393. Doi:10.1136/bmjqs-2012-000951

Pedhazur E. J. & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Pike, T. & O'Donnell, V. (2010). The impact of clinical simulation on learner self-efficacy in pre-registration nursing education, *Nurse Education Today, 30*, 405-410.

Polit, D. F. & Beck, C. T. (2010). *Essentials of nursing research: Appraising evidence for nursing practice* (7th ed.). Philadelphia: Lippincott Williams & Wilkins

Porter,C. O. L. H., Gogus, C. I., & Yu, R. C. (2011). The influence of early efficacy beliefs on teams' reactions to failing to reach performance goals. *Applied*

Psychology: An International Review,*60*(4), 645-669. Doi 10.1111/j.1464-0597.2011.00451.x

Portney, L. G. & Watkins, M. P. (2009). *Foundations of clinical research: Applications to practice* (3rd ed.). Upper Saddle River, NJ: Pearson Education.

Prince, C. & Salas, E. (1993). Training and research for teamwork in the military aircrew. In E. L. Weiner, B. G. Kanki & R. L. elmreich (Eds.), *Cockpit resource management* (337-366). Orlando, FL: Academic.

Quality and Safety Education for Nurses (2012). Quality and safety teamwork and collaboration competencies. Retrieved from http://www.qsen.org/ksas_prelicensure.php#teamwork_collaboration

Riley, W., Davis, S., Miller, K., Hansen, H., Sainfort, F., & Sweet, R. (2011). Didactic and simulation nontechnical skills team training to improve perinatal patient outcomes in a community hospital. *The Joint Commission Journal on Quality and Patient Safety*,*37*(8), 357-364.

Robertson, D. S. & Felicilda-Reynaldo, R. F. D. (2015). Evaluation of graduate nursing students' information literacy self-efficacy and applied skills. *Journal of Nursing Education*, *54*(3), S26-S30. Doi 10.3928/01484834-20150218-03.

Rosen, M. A., Salas, E., Wilson, K. A., King, H. B., Salisbury, M., Augenstein, J. S. Robinson, D. W., & Birnbach, D. J. (2008). Measuring team performace in simulation-based training: Adopting best practices for healthcare. *Simulation in Healthcare*, *3*(1), 33-41, doi: 10.1097/SIH.0b013e3181626276.

Rosen, M. A., Weaver, S. J., Lazzara, E. H., Salas, E., Wu, T., Silvestri, S., Schiebel, N., Almeida, S., & King, H. B. (2010). Tools for evaluating team performance in

simulation-based training. *Journal of Emergencies, Trauma, and Shock,3*(4), 353-359. doi:10.4103/0974-2700.70746.

Smith, E.L., Cronenwett, L., & Sherwood, G. (2007). Current assessments of quality and safety education in nursing. *Nursing Outlook, 55*(3), 132-137.

Sullivan, D. T., Hirst, D., & Cronenwett, L. (2009). Assessing quality and safety competencies of graduating prelicensure nursing subjects. *Nursing Outlook, 57,* 323-331. Doi:10.1016/j.outlook.2009.08.004

The Joint Commission (2008). Behaviors that undermine a culture of safety, Sentinel Event Alert, Issue 40. Retrieved from http://www.jointcommission.org/sentinel_event_alert_issue_40_behaviors_that_undermi ne_a_culture_of_safety/default.aspx)

Thomas, E. J., Sherwood, G. D., & Helmreich, R. L. (2003). Lessons from aviation: Teamwork to improve patient safety. *Nursing Economics, 21*(5), 241-243.

Thomas, E. J., Williams, A. L., Reichman, E. F., Lasky, R. E., Crandell, S., & Taggart, W. R. (2010). Team training in the neonatal resuscitation program for interns: Teamwork and quality of resuscitations. *Pediatrics, 125*(3), 539-546. Doi: 10.1542/peds2009-1635

Trochim, W. M. (2006). *The research methods knowledge base*, (3rd ed.). Web Centre for Social Research Methods, viewed 7 May 2015, <http://socialresearchmethods.net/kb/htm>.

Zhang, C. (2013). E*ffects of simulation-based interprofessional education on teamwork.* (Unpublished Doctoral Dissertation). University of Nebraska Medical Center, Omaha, NE.

Zhang, C., Miller, C., Volkman, K., Meza, J., & Jones, K. (2015). Evaluation of the team performance observation tool with targeted behavioral markers in simulation-based interprofessional education. *Journal of Interprofessional Care, 29*(3), 202-208. Doi: 10.3109/13561820.2014.982789

LeAnn J. Chisholm PhD(c), RN, CEN
P.O. Box 10081, Beaumont, TX 77710
(409) 880-8862
Leann.chisholm@lamar.edu

EDUCATION

University of Texas Medical Branch, Galveston, Texas
(PhD education in progress) **[Anticipated Aug. 2016]**
Dissertation: Validating Targeted Behavioral Markers for Teamwork Skill and Performance Outcomes in Simulation

Lamar University, Beaumont, TX **[2004]**
Master of Science in Nursing, Nursing Education
Thesis: "Nursing Student Motivating Factors in the Clinical Setting"

Lamar University, Beaumont, TX **[1997]**
Bachelor of Science in Nursing

ACADEMIC AWARDS

Texas Organization of Baccalaureate and Graduate Nursing Education Programs Excellence in Research Grant/Award	**2016**
Faculty Daisy Award	**2016**
UTMB School of Nursing Salute to Nursing Scholars Academic Achievement Award	**2015**
Robert Bennett Tuition Scholarship	**2013**
Regina R. & Alfonso J. Mercatante Memorial Scholarship	**2013**
Texas Organization of Baccalaureate and Graduate Nursing Education Programs Innovation in Nursing Education Award	**2012, 2011, 2010**
Daisy Award for Extraordinary Nurses	**2009**
Faculty Daisy Award, Nominee	**2012, 2011**
Dishman Faculty Excellence Award	**2005**
Army Nurse Corps Spirit of Nursing Award	**1997**
Lamar University Texas Nursing Student Association (LUTNSA) Award for Leadership	**1997**
Tim Halley Memorial Scholarship	**1997**
Maimie McFaddin-Ward Scholarship	**1997**

TEACHING EXPERIENCE

Lamar University, Beaumont, TX
Instructor
Revised basic clinical skills courses to utilize scenarios and promote clinical context. Continued promotion of best practices and evidence-base for clinical nursing skills. **2014 - current**

Developed and taught two online RN-BSN courses (Comprehensive Health Assessment and Compromised Adult Health States). **2012 - 2014**

Advanced Medical-surgical courses at Baccalaureate and Associate Degree Levels, Synthesis Seminar, Preceptorship Synthesis Experience **2004-present**

2004-present

Developed syllabus and overall course structure, including weekly lab practicum, and administered all grades.	**2005-present**
Collaborated on curriculum and exam development, met with students upon request, and graded all written work, including written papers.	**2004 – current**
Conduct Academic Advising for baccalaureate students from first to final semester in the nursing program.	**2004 - 2014**
Supervised students in medical surgical, critical care, and long-term acute care settings	

ACADEMIC ADMINISTRATIVE & RELATED HIGHER EDUCATION EXPERIENCE

Lamar University, Beaumont, TX **Director, Edna Horn Gay Learning Center –** Provide for a quality learning environment to enhance teaching/learning activities within the Learning Center and Dishman Department of Nursing. Administers all simulation labs and personnel including basic first level skills, advanced nursing skills, high-fidelity simulation, student workers, lab assistants, and serves as a resource for simulation pedagogy. Lead faculty and staff in the development, implementation, and evaluation of simulation learning activities. Collaborate with faculty to evaluate current equipment and media needed to achieve student learning outcomes and program outcomes.	**July 2014 - current**
High-fidelity Simulation Lab Coordinator – Promote excellence in the Dishman Department of Nursing through the administration of evidence-based high-tech nursing simulation scenarios which promote critical thinking and patient safety. Facilitate development of simulations, lead in review of simulations, develop guidelines and policies related to high-fidelity simulations, conduct faculty development and simulation technology competence, promote evidence-based use of simulated patients across the nursing curriculum, utilize volunteer nursing students to assist with underclassmen simulation experiences. Worked with a team to develop a new Simulated Patient Program which is integrated across the curriculum.	**2006 – Present**
Lamar University, Beaumont, TX **Dishman Department of Nursing Simulation Committee Chair** Facilitate simulation development and evaluation based on current best practices.	**2012 - Present**
Lamar University, Beaumont, TX **Dishman Department of Nursing Curriculum Co-chair** *Facilitate rigorous and systematic curriculum review in the Dishman Department of Nursing resulting in continued Accreditation from the National League of Nursing Accreditation Commission*	**2009 – Present**
Lamar University, Beaumont, TX **RN-BSN Articulation Committee** *Developed Comprehensive Holistic Health Assessment (RNSG 4316) and Compromised Health States (NURS 4620) courses which promote scholarly learning for registered nurses pursuing a baccalaureate nursing degree. These courses integrate meaningful use of technology, evidence-based practice, standards of care, nursing theory, and baccalaureate level writing skills within innovative and student-friendly courses. Review and recommend strategies to promote rigor and achieve accreditation standards for RN-BSN courses.*	**2013 - 2015**
Lamar University, Beaumont, TX **Continuing Nursing Education Committee Member**	**2008-2011**

Work with a faculty team to develop and administer academic presentations to disseminate nursing knowledge.

Lamar University, Beaumont, TX
Student Faculty Council — **2010 - Present**

Works with Department Chair, Director of Undergraduate Nursing Studies, Committee Chairs, Coordinators, and student representatives to address issues of student and faculty concern. Addresses and promotes resolution of issues in a manner which promotes respect and concern for students.

Lamar University, Beaumont, TX
Continuing Nursing Education Committee Member — **2008 – 2011**

Works with Continuing Nursing Education Committee to develop and implement educational programs which meet national and local community nursing goals.

Lamar University, Beaumont, TX
Faculty Practice Committee Member

Work with faculty team to promote and develop policies conducive to Nurse Educators interested in maintaining specialty certifications and clinical practice expertise.

CERTIFICATIONS

Drexel University, Philadelphia, PA
Certification in Simulation — **August, 2011 (does not expire)**

Certified Emergency Nurse (CEN), Board of Certification for Emergency Nursing (BCEN) — **2008 – current**

Alumnus CCRN, American Association of Critical Care Nurses (AACN) — **2007 – current**

Tulane University, New Orleans, LA
TeamSTEPPs® Master Trainer Course — **May 2015 (does not expire)**

Advanced Cardiac Life Support (ACLS) — **Current – 2017**

Basic Life Support (BLS) Instructor — **Current – Sept. 2016**

SIGNIFICANT PROFESSIONAL PUBLICATIONS

- Curl, E. D., Smith, S., **Chisholm, L**. A., McGee, L. A., & Das, K. (2016). Effectiveness of Integrated Simulation and Clinical Experiences Compared to Traditional Clinical Experiences for Nursing Students. *Nursing Education Perspectives*, *37*(2), 72-77 6p. doi:10.5480/15-1647

- **Chisholm, L**, Moss, P., Hale, G., Cochran, G., Goodwin, M. & Rivers, D. (2014). *Transforming the healthcare environment through an interprofessional and intraprofessional disaster simulation in order to improve communication and teamwork*. Abstract published in the proceedings for Transformation: Health Care Strategies Annual Research Day, Sigma Theta Tau International, Kappa Kappa Chapter, Beaumont, Texas

- **Chisholm, L**., Pipkins, C., Curl, E.D., Heinz, C., & Moore, M. (2014). *"Talk to a mannequin" or "interact with a patient": Advantages of using simulated patients.* [Abstract]. *Fifteenth National Conference of the Society for the Advancement of Modeling and Role-Modeling,* Erlinger, KY.

- Williams, S. & **Chisholm , L**. (2013). *Designing the Integration and Evaluation of a Faculty Developed Electronic Health Record and Web Portal. Abstract published in the proceedings for the National League for Nursing Technology Conference, Philadelphia, PA.*

- Williams, S. & **Chisholm, L.** (2012). *Integration of Informatics through an intranet hospital site and electronic health record. Abstract published in the proceedings for the American Association of Colleges of Nursing 2012 Baccalaureate Education Conference, San Antonio, TX.*

- Goodwin, M., Kilgore, R., Harding, R., & **Chisholm, L**. (2012). Using Simulation to Add Context to Concept Application for Leadership and Management Baccalaureate Nursing Students [Abstract]. 11th Annual INACSL Conference, San Antonio, TX.

- **Chisholm, L.** J., Curl, E. D., & Smith, S. (2011). Acute respiratory distress simulation for advanced medical-surgical/ critical care nursing courses. F.A. Davis.

- **Chisholm, L.J**., Hale, G., Williams, S., Jones, T., Welch, A., Walden, G., McKinley, C. and Stephens, S. (2009, April). *Interdisciplinary collaboration among undergraduate healthcare providers: A critical link in developing the art of holistic care in a high-tech healthcare environment.* Abstract published in the proceedings for the 27th Annual International Nursing Technology Conference, Washington D.C.

- **Chisholm, L.J.,** Hale, G., Williams, S., Jones, T., Welch, A., Walden, G., McKinley, C. and Stephens, S. (2009, April). *Empowering undergraduate healthcare providers: Critical lessons in interdisciplinary communication and end of life care.* Abstract published in the proceedings for the Geriatric Nursing, Education and Simulation Conference, Durham, NC.

- Curl, E. D., **Chisholm, L**., Smith, S., Hamilton, J., & McGee, L. (2007, February). *Best practice for simulation module development*. Abstract published in proceedings for the International Nursing, Computer and Technology Conference, San Francisco, CA.

- Curl, E., Smith, S., **Chisholm, L**., Hamilton, J., & McGee, L. (2007). *Collaborative approaches for developing evidence-based simulation modules.* Paper presentation at 18th International Nursing Research Congress, Sigma Theta Tau International, Vienna, Austria.

- Curl, E. D., Smith, S., **Chisholm, L.**, Hamilton, J., & McGee, L. (2006, October). *Developing evidence-based simulation modules through collaboration.* Abstract published in proceedings for the Professional Nurse Educators Group, Burlington, VT.

- Curl, E. D., Smith, S., **Chisholm, L.**, Hamilton, J., & McGee, L. (2007). Multidimensional approaches to extending nurse faculty resources without testing faculty's patience. *Journal of Nursing Education 46*(4), 193-195.

- Kappa Kappa Quarterly Review (2006-2010). **Chisholm, L**. (Ed.). 1(1), 1-4.

- **Chisholm, L**. (2005). Nursing student motivating factors in the clinical setting, Sigma Theta Tau International Kappa Kappa Chapter Annual Research Day. Beaumont: Lamar University.

- **Chisholm, L.** (2004). Nursing student motivating factors in the clinical setting. Unpublished master's thesis, Lamar University, Beaumont, Texas, USA.

BOOK CHAPTER REVIEWS

Reviewed Renal Chapter in Morton, P.G. & Fontaine, D.K. (2012). *Critical care nursing: A holistic approach* (10th ed.). Philadelphia, PA: Lippincott, Williams, & Wilkins.