EXPLORING FACULTY PREPARATION FOR AND USE OF DEBRIEFING WITH THE DEBRIEFING FOR MEANINGFUL LEARNING INVENTORY

Cynthia Diane Bradley

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

September, 2016

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

Doctoral Committee

Kristina Thomas Dreifuerst, PhD, RN, CNE, ANEF, Chair

Pamela M. Ironside, PhD, RN, FAAN, ANEF

Amy Hagedorn Wonder PhD, RN

Barbara Manz Friesth, PhD, RN

Pamela R. Jeffries, DNS, RN, FAAN, ANEF

July 19, 2016

Dedication

This work is dedicated to my friends, family, and colleagues who have provided their unending support through the journey of my doctoral program. I am forever indebted to each of you who never ceased to offer a listening ear, a shoulder to cry on, and unlimited words of encouragement. The personal challenges I encountered on this path presented not only opportunities for my own growth, but stepping stones for new levels of relationship with each of you. The past three years were filled with many personal and professional life changes, and I am indebted to each who played significant roles in supporting me through these ups and downs.

I dedicate this work to the memory of my mother, who saw only the beginning of this doctoral journey. Her example of faith, courage, and persistence are qualities that I possess as a result of the love and life she poured into me. I dedicate this also to my father, who has remained an example of steadfast faith and love throughout my life. His example of devotion to family, and commitment to pursuing education impacted my life in ways I never understood until recent years. Perhaps it was my typing of his thesis on an electric typewriter that planted a seed that culminated in my own dissertation so many years later.

I also dedicate this work to the future of healthcare education. May we forever explore, learn, and grow.

Acknowledgements

Truly, it takes a village to support the work of a dissertation. There are many to whom I am indebted for the successful completion of this work. First, I am grateful for each of the members of my dissertation committee. Dr. Pamela Ironside, I cannot even begin to express my gratitude for how much I have learned from you as a mentor, guide, and role model. Dr. Amy Wonder, I have many treasured memories of both intellectual and emotional conversations that guided me through trying times, as well as encouraging my progress. Dr. Barbara Freisth, I am thankful for your willingness to come alongside this process and contribute careful attention to detail and guiding comments. Dr. Pamela Jeffries, I am grateful for your presence and your wise counsel. From our first conversation, you pushed me to think beyond what I believed I was capable of and continually guided me to consider challenging perspectives.

And finally, Dr. Kristina Dreifuerst, my chair and trusted mentor. Where do I begin? Never did I dream this journey would unfold like it has, beginning with the first day we met. I believe we innately knew there would be incredible things in store for the future, both personally and professionally. Despite the fact you have read millions of words I have written over the past three years, I am now struggling to describe all you have been to me. On the days I thought I literally could not take one more step, you found a way to boost and buoy my feeble efforts and keep me moving forward. I am forever grateful to you.

I am blessed with many friends and family who have supported me in so many ways. My sons Nick, Jordan, Michael, and Matthew kept me grounded by their living examples of what is most important in life. Thank you to each of my colleagues who supported me through listening to my incessant chattering about my progress. I am thankful for my many friends who cheered me on and motivated me to keep going, even when I thought I could not. Deb, how can I possibly express how your kindred spirit and friendship has propelled me along this path. Barb, cheese and wine literally transported our friendship through dark places even when we both thought we would implode from our dissertation whining. Sara, your listening ear and caring spirit cheered me on numerous times. Marcie, you never shied away from tearful conversations and offered continual support. Jan, Teresa, and Ann, thank you for always believing in me.

I am thankful for the support of the National League for Nursing and the Jonas Center for the NLN Jonas Scholars program, the Indiana University School of Nursing for the Florence Nightingale Scholarship and the Emily Holmquist Award, and Dr. Mary Anne Rizzolo for the NLN Doctoral Research Award. This dissertation research would not have been possible without this funding and support.

Finally, I am grateful for those who contributed to the completion of this work. Dr. Darrell Spurlock, thank you for your inspiration, and your direction in the scale development process. Thank you to the expert raters and debriefers who contributed to scale development and pilot testing. I am indebted to the debriefers who volunteered their time to describe their debriefing training for this study. This work would not have been possible without your valuable input.

Cynthia Diane Bradley

EXPLORING FACULTY PREPARATION FOR AND USE OF DEBRIEFING WITH THE DEBRIEFING FOR MEANINGFUL LEARNING INVENTORY

Debriefing is the most significant component of simulation, yet the impact of debriefer training for this critical time of learning has not been reported. Although training of debriefers in the use of a structured debriefing method has been recommended by nursing organizations and regulating bodies, a description of the impact of training on the understanding and application of debriefing within nursing programs is largely unknown.

Debriefing for Meaningful Learning© (DML) is a structured, evidence-based debriefing method that promotes the development of clinical reasoning among prelicensure nursing students. DML has been adopted for use across the curriculum throughout nursing education. However, little is known regarding how debriefers are trained in this method and how that training impacts their understanding and application of the method.

The DML Inventory (DMLI), was developed and tested for this research study, and used to assess and document 234 debriefers' understanding of the central concepts of DML and subsequent application of DML behaviors during simulation debriefing with prelicensure baccalaureate nursing students. Statistically significant differences were found between those debriefers who had and had not been trained in DML. Statistically significant differences were also found in the understanding of the DML central concepts, and in the application of DML based on the types of training the debriefers received.

The data indicate that DML trained debriefers consistently apply more DML behaviors than those who had not received training, and that multiple sources of training resulted in a more consistent application of DML debriefing behaviors. Furthermore, understanding the central concepts of DML resulted in debriefers consistently applying more debriefing behaviors consistent with the DML design.

This study contributes to the growing body of knowledge of debriefing in nursing education and training in evidence-based debriefing methods, by providing a tested instrument that can be used to assess debriefers using DML. The findings also demonstrate the impact of the type of training on how structured debriefing methods are then implemented in teaching-learning environments, which can be used to improve future training.

Kristina Thomas Dreifuerst, PhD, RN, CNE, ANEF, Chair

Table of Contents

List of Tables

List of Figures

Abbreviations

Abbreviations	Terms
CCU	Central Concept of Understanding
DML	Debriefing for Meaningful Learning©
DMLES	Debriefing for Meaningful Learning Evaluation Scale
DMLI	Debriefing for Meaningful Learning Inventory
LCFA	Latent Class Factor Analysis
NCSBN	National Council of State Boards of Nursing
NLN	National League for Nursing
NSS	National Simulation Study

Chapter I Introduction

Background of the Study

There is broad consensus across higher education that the preparation to teach within a college or university is inadequate (Eckert, 2014; Gurung & Schwartz, 2009; Malott, Hall, Sheely-Moore, Krell, & Cardaciotto, 2014; Monk, 2015; McAllister, Cantrell, & Meakim, 2014; Oleson & Hora, 2014; Sawatzky & Ennis, 2009). Traditionally, content expertise has been the most respected feature of a university professor across disciplines (Cangelosi, Crocker, & Sorrell, 2009; Dent & Harden, 2013; Postareff, Lindblom-Ylanne, & Nevgi, 2007). Within the health professions, it is common for a practitioner to assume an academic role to teach content specific to the profession, regardless of whether they have received training to teach (McLean, 2010). Gurung and Schwartz (2009) identified that while it would not be assumed that a teacher intuitively knows the content to teach, it is expected that the ability to teach discipline-specific content is intuitive. The practice of teaching, however, is not a natural derivative of content expertise, but requires training in evidence-based pedagogy and specific teaching methods (Harden & Crosby, 2000).

As the learner-centered model expands and transforms within higher education, the need for excellence in teaching has also progressively emerged (Fincher & Work, 2007; Fitzmaurice, 2010; Kalb, 2008). The literature within higher education reflects increasing attention on testing and implementing a teaching practice that is evidence-based (Bok, 2009; Groccia & Buskitt, 2011; Ferguson & Day, 2005; Ironside & Valiga, 2008; Kalb, O’Conner-Von, Brockway, Pierson, & Sendelbach, 2015; McAllister, Oprescu, & Jones, 2014; Patterson, 2009), rather than a practice that is dependent on teaching as one was taught (Bartels, 2007; Cangelosi et al.; Gardner, 2014; Ja'Afar, 2012; Shulman, 2005). As opposed to relying on experiential and anecdotal evidence, higher education has gradually embraced evidence-based teaching as a foundation for practice (Groccia & Buskitt, 2011).

One teaching-learning method that is supportive of how people best learn is debriefing. Debriefing is an intentional time at the end of a learning experience for a debriefer to engage learners in a collaborative, guided, reflective discussion in order to make meaning of the experience (Cantrell, 2008; Cheng, Rodgers, van der Jagt, & O'Donnell, 2012; Dreifuerst, 2009; Fanning & Gaba, 2007; Pearson & Smith, 1986). During debriefing, the debriefer and learners revisit a simulation or clinical experience, make sense of what occurred, correct learner misinterpretations and incorrect actions, and acknowledge and reinforce appropriate thinking and actions (Dreifuerst, 2009; Fanning & Gaba, 2007).

Simulation literature offers increasing evidence of improved learner outcomes through debriefing (Fanning & Gaba, 2007; McGaghie, Issenberg, Petrusa, & Scalese, 2010; Hayden, Smiley, Alexander, Kardong-Edgren, & Jeffries, 2014). An evidence-based debriefing provides both debriefers and learners with an intentional, systematic process leading to strong learner outcomes (Dreifuerst, 2012; Mariani, Cantrell, & Meakim, 2014). Although an evidence-based teaching method may have demonstrated positive outcomes in rigorous research studies, using that method in a teaching-learning environment does not guarantee consistent learning outcomes if the debriefer lacks understanding of its significant concepts, or if the method is not applied as it was designed and tested. The literature is lacking in a description of the training a debriefer needs to consistently apply a debriefing method in its original design (Garden, LeFevre, Waddington, & Weller, 2015; Waznonis, 2015). Additionally, it remains unknown how training of the debriefer in evidence-based methods is being translated into teaching practice, and what this effect is on learner outcomes.

Debriefing for Meaningful Learning© (DML) is an evidence-based debriefing method that promotes the development of clinical reasoning, or thinking like a nurse, among prelicensure nursing students (Dreifuerst, 2012; Forneris et al., 2015; Mariani, Cantrell, Meakim, Prieto, & Dreifuerst, 2013). DML is used to debrief students in simulation and other clinical learning environments to cultivate reflective thinking and enhance learning (Dreifuerst, 2015). DML was also the debriefing method used for the National Council State Boards of Nursing (NCSBN) National Simulation Study (NSS) that demonstrated positive student learning outcomes (Hayden et al., 2014).

While DML has been widely adopted in nursing programs, it is not known how debriefers have received training in DML, nor how consistently the central concepts of DML are understood or applied (K.T. Dreifuerst, personal communication, November 6, 2014). The central concepts of DML include reflection-in-action, thinking like a nurse, reflection-on-action, challenging taken-for-granted assumptions, and reflection-beyond-action (Dreifuerst, 2012). These concepts are highly complex constructs, requiring specific training to ensure that debriefers gain understanding to apply them when during debriefing. Without training in DML, it cannot be assumed that a debriefer would ensure consistent learning outcomes when attempting to use this method (Jeffries, Dreifuerst, Kardong-Edgren, & Hayden, 2015). Yet, this relationship between the training of debriefers and the resulting understanding and application of DML is unexplored.

In addition, it is not known how DML is applied in debriefing, regardless of whether the debriefer was trained or not trained. Although debriefers may identify with using DML, it is possible that no training was received, limiting the ability to apply this method of debriefing. It is also possible that training was received, but application is limited due to lack of retention, or even intentionally choosing to vary the method from its design and purpose. Moreover, some debriefers exercise liberty with incorporating additional teaching and debriefing constructs into how they debrief students, regardless of the method. Without further exploration, there are many unknowns regarding the impact of training on debriefing application, which could result in debriefers claiming to use an evidence-based method when the reality is that they actually are not.

Although some constructs of debriefing have been studied (Fey, 2014; Waznonis, 2015), there are no reports of a psychometrically tested instrument measuring a debriefer's ability to adhere to the behaviors consistent with a specific debriefing method such as DML. Therefore, a formative, evaluative, behaviorally anchored scale, the Debriefing for Meaningful Learning Evaluation Scale (DMLES) was developed based on the framework of DML to evaluate the ability of a debriefer to implement this method (Bradley & Dreifuerst, 2016). Findings from the pilot testing revealed that the DMLES is a valid measure of DML debriefing behaviors in the context of simulation debriefing with prelicensure nursing students. These findings supported the further investigation of debriefing practices by debriefers who debrief prelicensure baccalaureate nursing students; particularly those who report they were trained in or use DML.

Statement of the Problem and Research Questions

Despite widespread adoption of DML across nursing education, little is known regarding how this debriefing method is being used during simulation debriefing with prelicensure baccalaureate nursing students. It is also unknown how various types and sources of DML training impact a debriefer's understanding of the central concepts of this debriefing method, and subsequent application of the method. The research questions guiding this study were as follows:

1. Is the DMLI a valid measure of DML understanding and application?
2. Is there a difference in how many of the central concepts associated with DML that debriefers understood, when they were grouped according to the training they received?
3. Is there a difference in how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received?
4. Is there an interaction between how many of the central concepts associated with DML that debriefers understood, and how many behaviors

associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received?

Purpose of the Study

The purpose of this study was to describe the impact of DML training on how consistently debriefers understand and apply the central concepts of DML, when grouped according to the type of training they received. While DML has demonstrated improvement in the clinical reasoning of prelicensure nursing students (Dreifuerst, 2012; Forneris et al., 2015), it is unknown how debriefers are trained in the method, or how this training impacts a debriefer's understanding and application of the method while debriefing prelicensure baccalaureate nursing students. Two groups of debriefers were compared and described according to the type of debriefing training received: debriefers who received DML training, and debriefers who received training in a method other than DML.

Significance of the Study

This study makes a significant contribution to advancing the science of nursing education research by describing the impact of training in an evidence-based debriefing method on faculty debriefer application with prelicensure baccalaureate nursing students. This research also contributes a new valid instrument designed to assess how debriefers debrief with the DML method when compared to its original design.

There is wide consensus that debriefing is the most significant component for student learning in simulation (Shinnick, Woo, Horwich, & Steadman, 2011). Because of this, regulations and guidelines were developed by the International Nursing Association for Clinical Simulation and Learning (INACSL Standards of Best Practice: Simulation[SM], 2013), the National League for Nursing (NLN, 2015), and the NCSBN (Alexander et al., 2015) that emphasize the importance of preparing debriefers with formal training in debriefing. However, little is known regarding how best to train debriefers to apply an evidence-based debriefing method. It is also not known how consistently DML is understood and subsequently applied by debriefers after training, since there has been no tool that specifically measures how the method is being used in debriefing.

This study contributes further disciplinary understanding of the impact of the types and sources of training debriefers commonly receive when learning how to implement a debriefing method. Conferences and workshops are reported as a primary source of training within nursing education (Patterson, 2009), and specifically in simulation debriefing in nursing education (Fey, 2014; Waznonis, 2015). Other reported sources of teacher training in nursing education include reading literature and learning from colleagues (Patterson, 2009). A description and comparison of the impact of each source of training is important to understanding how nurse educators best learn to debrief, and retain debriefing knowledge, for designing future faculty development. This understanding also has implications for program quality improvement, and faculty development in regards to cost, time, and allocation of resources.

Definition of Terms

The following terms will be used throughout this study. For the purpose of this study, these terms are defined as follows.

Assessment

Assessment in education is the gathering of information about students, faculty, curricula, programs, or institutions (Astin, 2012). Assessment is used to provide information about learning needs, and to improve teaching and learning (Billings & Halstead, 2013).

Debriefer

An individual who facilitates a debriefing after a learning experience with students, through engagement in a collaborative, reflective discussion. For the purposes of this study, it was assumed that debriefers were nursing faculty who facilitate simulation experiences including debriefing for baccalaureate nursing students.

Debriefing

Debriefing is a teaching-learning method used by a debriefer to engage learners in a collaborative, reflective discussion following a simulation experience (Meakim et al., 2013). Through this learner-centered conversation, debriefing deepens the understanding of learners by integrating experience with reflective thinking (Decker et al., 2013; Dreifuerst, 2009; Shinnick et al., 2011).

Debriefing for Meaningful Learning (DML)

DML is a structured debriefing method that uses Socratic questioning and guided reflection to cultivate a student's reflective thinking and promote meaningful learning (Dreifuerst, 2012). A debriefer guides students through reflection-in-action, reflection-on-action, and reflection-beyond-action using the six E's of debriefing: engage, explore, explain, elaborate, evaluate, and extend.

Debriefing for Meaningful Learning Evaluation Scale (DMLES)

The Debriefing for Meaningful Learning Evaluation Scale (DMLES) is a 31-item rating scale behaviorally anchored in the iterative process of DML (Bradley & Dreifuerst, 2016). The observable behaviors of each of the six E's (engage, explore, explain, elaborate, evaluate and extend) are structured as 31 distinct items to be scored with the binary options of 1) *present* or 2) *not present*.

Debriefing for Meaningful Learning Inventory (DMLI)

The Debriefing for Meaningful Learning Inventory (DMLI) is a 57-item subjective measure of debriefing behaviors consistent with DML. Items one through 51 describe behaviors consistent with DML, and are scored with ordinal frequency options: *always*, *sometimes*, *never*. Items 53 through 57 describe the central concepts of the method, and are scored with binary options of *yes* or *no*.

DML Group

For this study, the DML group is defined as the group of debriefers who identify they have received training in DML, and use the method in simulation debriefing with baccalaureate nursing students.

Evaluation

Evaluation is a process of determining progress by making a judgment, with the intention of making improvements (Billings & Halstead, 2013).

Facilitator

A facilitator is an individual who provides guidance and support during simulation-based learning experiences (Meakim et al., 2013).

Mean DMLI sum

The mean DMLI sum is defined as the summative score of items one through 52 on the DMLI. This value represents the number of DML behaviors a debriefer consistently applies during debriefing.

Non-DML Group

For this study, the non-DML group is defined as the group of debriefers who identify they have received training in any other debriefing method other than DML, and use that method in simulation debriefing with baccalaureate nursing students.

Theoretical Framework

The Input-Environment-Output (I-E-O) Model is a conceptual framework used for assessment in higher education (Astin, 2012). The I-E-O Model was developed by Astin (1968), a clinical and counseling psychologist, as a developmental framework from which to view behavior. The I-E-O Model was deductively developed as a result of Astin's research in assessing doctoral productivity, specifically, on his findings that collecting data on inputs, outcomes, and the education environment is foundational to every educational assessment project (Thurmond & Popkess-Vawter, 2001). The purpose of the I-E-O model is to describe, explain, and predict behavior by describing assessments in higher education, explain how outcomes are influenced by input and environmental variables, and predict relationships between concepts as well as predicting outcomes resulting from those relationships (Cappell & Karmens, 2002).

The I-E-O Model was developed for use in natural settings, rather than in experimental settings (Astin, 1993). The model is useful in any social or behavioral science that observes human beings and the influence of the environment on their development (Thurmond & Popkess-Vawter, 2001). As a result, the I-E-O Model has been used as a framework for assessing the impact of simulation on student learning outcomes in nursing education (Prion, 2008), as well as in numerous research studies to assess the relationships between student inputs, environmental factors, and student outcomes (Astin, 1968; Astin & Sax, 1998; Campbell & Blakely, 1996; House, 1999; Kelly, 1996; Thurmond & Popkess-Vawter, 2001).

Astin (1993) describes inputs as the qualities that a student brings to an education program. These qualities can include demographic information, performance tests, prior knowledge, educational background, behavior pattern, motivation, or life goals. Environment refers to the experiences of a student during the educational program, and includes anything that may impact a student such as the instructor, curricula, facilities, climate, teaching style, peers, or organizational affiliation (Cappell & Karmens, 2002). Output is described as “the talents we are trying to develop in…educational programs” (Astin, 1993, p.18). This includes outcome variables such as posttests, consequences, exam scores, performance scores, degree completion, and overall satisfaction (Pike & Killian, 2001).

As a framework for guiding the assessment of students, the I-O-E Model underpins this study because when receiving training, a debriefer assumes the role of a student. The focus of this study is the training a debriefer receives in DML. In alignment with the I-E-O Model, the debriefer brings individual student inputs at the time of DML training. Descriptions of many of these inputs will be collected in this study, including demographic characteristics, previous education, previous teaching experience, and previous simulation experience. Although not a focus of this study, there are other inputs that may influence DML training including motivation, life goals, behavioral patterns, and aptitude.

In this study, environment is the described characteristics of the DML training. This includes the type of training, the training source, the trainer, the organizational affiliation of the trainer, the length of the training session, and the nature of an evaluation component of the training session, if evaluation was included in the training. Other environmental variables that may also influence the training include facilities, instructor, climate, teaching and learning styles, and the absence or presence of peers.

Output in this study is the understanding and application of DML in debriefing. Astin (1993) described outputs as the talents to be developed; the purpose of DML training is to prepare debriefers to understand the central concept of the method, and apply this in debriefing. For this study, the output will be measured with the DMLI through self-report of a debriefer describing how they understand and apply DML in a debriefing with prelicensure nursing students.

Organization of the Study

This research study will be presented in five chapters. Chapter I consists of the background of the study, problem statement, purpose, significance of the study, definition of terms, theoretical framework, and research questions of the study. Chapter II presents a review of the literature including debriefing, training the debriefer, retention, DML, and evaluation of the debriefer. Chapter III presents the methodology used for this research study. This includes selection of participants, instrumentation, data collection, and data analysis procedures. Chapter IV presents the findings of the research including demographic information, confirmatory factor analysis, and results of the data analyses for the research questions. Chapter V provides a summary of the study, a discussion of the findings with implications for nursing education, recommendations for further research, and conclusions.

Chapter II Review of the Literature

This chapter presents literature pertinent to the study of the impact of training in Debriefing for Meaningful Learning© (DML) on how consistently debriefers trained in this method understand and apply its central concepts during simulation debriefing with prelicensure baccalaureate nursing students. DML is an evidence-based debriefing method that has demonstrated a positive effect on the development of clinical reasoning among prelicensure nursing students (Dreifuerst, 2012). Little is known regarding the application of evidence-based debriefing methods, or the impact of common types and sources of training on how these methods are applied. This chapter includes a review of the literature related to debriefing, training the debriefer, knowledge retention, DML, and evaluation of the debriefer.

Debriefing

There is wide agreement that debriefing is the most significant component of a simulation experience, due to the sense-making and depth of learning that occurs during this planned time after a learning experience (Cantrell, 2008; Decker et al., 2013; Dreifuerst, 2009; Dreifuerst & Decker, 2012; Fanning & Gaba, 2007; Mariani, Cantrell, & Meakim, 2014; Shinnick, Woo, Horwich, & Steadman, 2011). Debriefing is a teaching-learning method used by a debriefer to engage learners in a collaborative, reflective discussion following a learning experience (Cantrell, 2008; Cheng, Rodgers, van der Jagt, Eppich, O'Donnell, 2012; Fanning & Gaba, 2007; Pearson & Smith, 1986). Debriefing provides the time and environment to reconstruct events within a group to facilitate shared learning (Fanning & Gaba, 2007).

Debriefing can be a type of formative assessment that transpires through an interaction between a debriefer and a group of learners, as the thinking and actions of the learning experience are collaboratively reexamined (Rudolph, Simon, Raemer, & Eppich, 2008). As such, debriefing involves active participation between a debriefer and learners. The debriefer intentionally cultivates open, trustful, and confidential communication to engage learners. This engagement facilitates the uncovering and reconstruction of thinking, decision, and actions that occurred during the learning experience (Dreifuerst, 2009; Neill & Wotton, 2011). As learners engage in experiential learning activities, making the connection between doing and thinking is important; it was this making sense of an experienced event that led to the concept of debriefing (Fanning & Gaba, 2007).

The ability to debrief learners is as important in simulation learning experiences as the ability to develop scenarios and use simulation equipment (Jeffries, 2005). Debriefing has been identified as necessary for consolidating and transferring learning (Rudolph et al., 2008). Because of this, ineffective debriefing skills could result in the learner not recognizing faulty assumptions, misconceptions, errors, and mismatched thinking and actions; this could lead to the learner's transfer of misconceptions into nursing practice (Dreifuerst, 2009).

History of Debriefing

Debriefing originated as a post-mission military experience in which participants collaboratively described the details of a mission to the military leaders who were not located at the front line of the conflict in order to reexamine the actions that occurred, and to develop strategies for future missions (Pearson & Smith, 1986). Following this format, the aviation industry engaged in simulation and debriefing as a result of accidents of highly skilled pilots in the 1970s which exposed the insufficiency of technical skills in ensuring positive outcomes (Billings & Reynard, 1984). Both novice and expert pilots then consistently practiced flight skills in simulated flight scenarios, followed by a facilitated debriefing in which they analyzed the scenario and their performance (Dismukes, Gaba, & Howard, 2006). It was within this context that educators began to recognize debriefing as an essential component of simulation.

Lederman (1992) described debriefing as a post-experience analytic process, then later outlined a debriefing framework of cognitive assimilation. This structured method of finding meaning from the experience was later organized with seven debriefing elements including: a debriefer, participants, a learning experience, impact of the experience, recollection of the experience, reporting of the experience, and processing the experience (Lederman, 1992). Debriefing occurred in three stages described as reflection and analysis, individual meaning related to the experience, and generalization and application of lessons.

Pearson and Smith (1986) expounded on Lederman's work by identifying three stages of debriefing summarized by the questions "What happened? How did the participants feel? What does it mean?" (p. 72). In this manner of debriefing, participants were led to describe the experience in a shared, non-threatening environment. Participants were encouraged to describe their feelings and reactions, regardless of whether this felt threatening to participants. In the final stage of debriefing, participants generalized key learning points from the experience in order to find immediate relevant meaning.

Analysis of Debriefing

Although debriefing continued to spread in use as simulation-based education proliferated throughout the health professions, the concept of debriefing had yet to be defined or explained through formal analysis. In order to contribute to simulation-education best practices and guide the future of simulation research, Dreifuerst (2009) analyzed the concept of debriefing by identifying its essential components according to the Walker and Avant (2005) method of concept analysis.

The defining attributes of a concept are the attributes that are most frequently associated with the concept, and provide the most insight (Walker & Avant, 2011). Dreifuerst (2009) defined debriefing with the attributes of reflection, emotion and emotional release, reception to feedback, summative evaluation, and integration of the new knowledge through assimilation, accommodation, anticipation. These defining attributes all contribute to a significant learning experience in debriefing. A debriefer guides learners to reflect in order to reexamine the learning experience. Emotion and emotional responses can enhance learning or can inhibit learning; it is necessary for the debriefer to facilitate the expression of emotions in order for learners to safely frame the experience. Reception is facilitated by the debriefer in helping the learners maintain openness to receiving feedback during debriefing. Summative evaluation should be communicated to the learners if this is part of a debriefing experience, and requires the debriefer to maintain a confidential, respectful environment. Integration of the simulation with guided reflection through a conceptual framework occurs through assimilation of knowledge, skills, and attitudes, accommodation, and anticipation. Anticipation is closely related to reflection, and requires the ability to look forward or reflect beyond the current action.

Dreifuerst (2009) identified the antecedents of debriefing as the patient's story, the physiological processes of the patient in the scenario, and the learning objectives for the scenario. Consequences of debriefing are impacted by the presence and quality of the defining attributes; these consequences include varying changes in critical thinking, clinical decision making, and clinical judgment.

Empirical referents of a concept are the phenomena that demonstrate a concept has occurred (Walker & Avant, 2011). Dreifuerst (2009) identified the empirical referents of debriefing as unstructured debriefing, structured for critiquing debriefing, and structured for reflection debriefing. Unstructured debriefing lacks direction; the debriefer does not engage students in making connections from the simulation. Debriefing structured for critique is focused solely on performance; the debriefer focuses on behaviors, skills, and simple decisions made in the simulation. In a debriefing structured for reflection, the debriefer engages students to reflect on the challenges of patient care required in the simulation, and a collaborative guided discussion with feedback that promotes a deeper level of learning.

Debriefing in Nursing Education

As simulation learning experiences have multiplied in nursing education, the rigor of simulation pedagogy has also increased. In 2005, Jeffries published the Simulation Model for the purpose of guiding the design, implementation, and evaluation of simulation in nursing education. Debriefing was identified in this model as one characteristic of the simulation design, although sometimes overlooked (Jeffries, 2005). Debriefing was described as a group discussion of the “process, outcome, and application of the scenario to clinical practice” (Jeffries, 2005, p. 101). A second variation of the model was published by Jeffries and Rodgers (2007) as the Nursing Education Simulation Framework. This model was then expanded into the NLN Jeffries Simulation Theory (Jeffries, 2016).

In 2011, the state of the science of simulation in nursing education and the use of the Nursing Education Simulation Framework was examined by Jeffries and simulation thought leaders of the International Nursing Association for Clinical Simulation in Learning (INACSL) (Jeffries, 2015). This effort led to the development of the *INACSL Standards of Best Practice: Simulation*^{SM}, which was intended to provide the first set of standards for evaluating and improving simulation practices. The standards were revised in 2013, and two more standards were added in 2015.

Within this set of standards, one standard focused on the debriefing process, Standard VI, which articulates that a debriefing session is to follow all simulation learning experiences to ensure best learning outcomes (Decker et al., 2013). In addition, Standard VI states that debriefers must be subject matter experts who adhere to a structured debriefing method during all debriefing sessions. This adherence requires that debriefers receive formal training in a debriefing method, with competency assessment of their debriefing skills. Standard VI identifies that the debriefer should gain knowledge of debriefing best practices related to debriefing format and facilitation of reflective dialogue; participate in formal education of a structured debriefing method; validate competence of debriefing skills with an established instrument, experienced debriefers, and learners; and maintain competence of debriefing skills through practice in simulation education.

Training the Debriefer

Despite the significance of debriefing to simulation learning, and the recommendations for training in this critical teaching-learning method, there is little research investigating how best to train a debriefer to apply a debriefing method. In a review of nursing research on simulation debriefing, Neill and Wotton (2011) noted the need for nurse educators to master the practice of debriefing. In addition, Raemer and colleagues (2011) identified the lack of research exploring the impact of debriefing training programs on learner outcomes and the quality of debriefing, as well as the lack of research investigating the frequency of necessary retraining.

To address the lack of training in debriefing, INACSL, the National League for Nursing (NLN), and the National Council of State Boards of Nursing (NCSBN) each made recommendations outlining the specific needs in preparing debriefers to implement evidence-based debriefing methods. INACSL Standard VI recommends that debriefers receive training in a structured debriefing method through formal teaching, continuing education, or mentorship (Decker et al., 2013). In addition, Standard VI states that competence should be validated through an established instrument and through the feedback of learners and experienced debriefers. The NLN (2015) reinforced this recommendation that nurse educators receive formal training in a theory-based debriefing method, and engage in on-going assessment of debriefing competence. The NLN also recommended the integration of theory-based debriefing throughout nursing curriculum, because of the positive learning outcomes.

Most recently, following the seminal National Simulation Study (NSS) (Hayden et al., 2014) the NCSBN issued simulation guidelines to provide state boards of nursing guidance for regulation language that would be used by nursing programs (Alexander et al., 2015). The NCSBN specified that the program "utilizes a standardized method of debriefing…using a Socratic methodology" (Alexander et al., 2015, p. 41), and that the debriefing should follow evidence-based literature. While debriefing training was not specifically identified in the report, the need for preparing faculty for simulation was addressed broadly, and recommended that nursing programs attain preparation consistent with the INACSL Standards of Best Practice (2015).

Despite these recommendations, findings of two national surveys of nursing simulation debriefing practices reveal that actual debriefing practices are largely not evidence based (Fey, 2014; Waznonis, 2015). Fewer than half of simulation educator respondents reported facilitating structured debriefing sessions (Waznonis, 2015), and fewer than one-third of respondents reported using a theory or model based debriefing method (Fey, 2014).

Fey (2014) reported that less than half of debriefers received debriefing training (47.1%, n = 228/484) and an assessment of their competence was reported by only 19% of respondents (n = 91/484). Of designated simulation experts who received training, the reported sources of that training included workshop attendance in an academic setting (47.3%, n = 181/484), a simulation consultant (39.93%, n = 153/484), a simulation center (31.1%, n = 119/484), using online offerings from the NLN-Simulation Innovation Resource Center (SIRC) (28.7%, n = 110/484), graduate coursework (17%, n = 65/484), and by participating in the NLN Simulation Scholars program (3.4%, n = 13/484).

In a separate survey, Waznonis (2015) reported that 94% (n = 195/208) of survey respondents reported receiving training in debriefing. Of the respondents who reported receiving training, the type of debriefing training received was described as training from mentors (47%, n = 97/217), conference and workshop attendance (40%, n = 83/219), manikin company representatives (36%, n =

75/219), vendor meetings (28%, *n* = 58/219), and local or regional training offering attendance (27%, *n* = 56/219). Other sources of training reported by respondents were participating in webinars (11%, *n* = 23/219), attending an SSIH-CHSE program (7%, *n* = 15/219), participating in the Center for Medical Simulation (6%, *n* = 13/219), and graduate level certificate programs (3%, *n* = 7/219).

Reports of Debriefing Training

While survey respondents reported receiving training in debriefing methods (Fey, 2014; Waznonis, 2015), little is known regarding the nature of these training offerings. Within debriefing literature, three descriptions of debriefing training have been reported. These three reports describe a training session designed for Pediatric Advanced Life Support (PALS) participants (Cheng et al., 2012), a train-the-trainer workshop for surgical simulation-based teaching (Paige, Arora, Fernandez, & Seymour, 2015), and training sessions for the NCSBN multi-site, longitudinal NSS in prelicensure nursing education (Jeffries, Dreifuerst, Kardong-Edgren, & Hayden, 2015).

In the first of the three reports, Cheng and colleagues (2012) described the evolution of learning experiences for life support training concurrent with the advancement of simulation technology, specifically the shift from instructor-learner feedback to debriefing. The American Heart Association (AHA) endorsed debriefing in resuscitation courses because of the related improved learning outcomes. In 2009, the AHA introduced a two-hour online interactive debriefing module for resuscitation course instructors. A debriefing tool for instructors was also introduced to be used as a cognitive aid during the course to learn debriefing concepts. The AHA then collaborated with the Winter Institute for Simulation, Education, and Research (WISER) to develop the Structured and Supported Debriefing Gather, Analyze, and Summarize model. The purpose of this model was to develop a structured debriefing approach to be adapted to any debriefing situation, and also to offer an online training for debriefing in PALS and Advanced Cardiac Life Support (ACLS).

In the second training report, Paige and colleagues (2015) described a train-the-trainer workshop to teach evidence-based debriefing components for surgical simulation-based training; two 90-minute training sessions were developed based on the framework of Kolb's Experiential Learning Theory (ELT). The Objective Structured Assessment of Debriefing (OSAD), an eight-item debriefing tool, was used to teach debriefing skills in the first session during the observation and evaluation of both live and recorded debriefings. In the second session, participants incorporated the debriefing principles into a practice session of their debriefing techniques, and were provided feedback on their debriefing performance. Effectiveness of the training sessions was assessed using a pre-post-workshop Likert scale questionnaire, measuring self-efficacy of debriefing skills; seven of eight items demonstrated statistically significant increases in self-efficacy.

The third and most recent report of debriefing training, the training of debriefers for the NCSBN NSS was described by Jeffries, Dreifuerst, Kardong-Edgren, and Hayden (2015). In preparation for participation in the NSS, study team members attended three separate multi-day workshops during the 12 months preceding the study. As part of this preparation, study team members received training in Debriefing for Meaningful Learning© (DML) by the author of the method. Study team members learned to use the DML worksheets, and facilitate debriefing discussion using Socratic questions to guide students in reflection-in-action, reflection-on-action, and reflection-beyond action. These study team members observed an expert model demonstration of a DML debriefing by the method developer, then practiced using the debriefing method. A return demonstration was evaluated prior to study team members using DML with students in their respective nursing programs. In addition, study team members' use of DML in debriefing was periodically observed and evaluated with the Debriefing Assessment for Simulation in Healthcare-Rater Version© (DASH-RV; Simon, Raemer, & Rudolph, 2009).

Common Sources of Training

To gain understanding of common sources of evidence used to build teaching practice, Patterson (2009) conducted a survey of nurse educators. Respondents reported using three main sources of evidence in informing teaching practice: colleagues, literature, and professional conferences. Although this report was limited by a small sample size, these findings are consistent with the findings of surveys of debriefing practices within nursing education as reported by Fey (2014) and Waznonis (2015).

A common source of learning for nurse educators is the NLN Education Summit. The purpose of this annual conference is to provide opportunities for nurse educators to acquire new learning, share scholarship, and collaborate in new opportunities that support lifelong learning (NLN, 2016). Each year the NLN offers continuing education opportunities, and hosts speakers on topics related to preparing nurses for a diverse, complex healthcare environment. In addition, a wide variety of concurrent sessions are offered to attendees. In 2014, 1,250 nurse educators registered to attend the NLN Education Summit (E. Tagliareni, personal communication, May 26, 2016). This number increased to a registration of 1,375 in 2015.

Retention of Knowledge and Skills After Training

Fanning and Gaba (2007) reported that the estimated half-life of professional knowledge gained through formal education or training may be as little as two to two and one-half years. However, retention can be as little as six to 12 months when both formal knowledge and a set of skills are required, such as resuscitation skills (O'Steen, Kee, & Minick, 1996).

Sousa (2000) defined retention as a measure of how good material is remembered over time. Retention involves retrieval of memories; the success of retrieval is dependent on effective encoding (Bruning, Schraw, & Norby, 2010). Effective encoding involves making associations with prior knowledge that facilitates future retrieval. Therefore, retention can be considered as the extent to which one can retrieve information from long-term memory. The strength of retention is dependent on how deeply information is processed by the working memory, and subsequently held within long-term memory.

Retention of knowledge and skills may also be directly linked to how the information was presented as well as the dose of the presentation (Bruning et al., 2010). Educational interventions have been tested to promote retention of knowledge and skills, such as spacing the presentation of information and training. Raman, McLauglin, Violato, Rostom, Allard, and Coderre (2010) found that a dispersed curriculum among gastroenterology residents, or learning experiences presented at time intervals, promoted knowledge retention at three months' post-test compared to one comprehensive offering. In addition, Kerfoot, and Brotschi (2009) reported that spaced educational offerings of weekly clinical scenarios and questions improved knowledge retention in two of the four topics presented during a urology surgery course, by third year medical students (n = 115).

Oermann and colleagues (2011) found that with increased deliberate practice on a voice activated manikin designed for cardiopulmonary resuscitation, nursing students demonstrated higher compression rates and statistically significantly better depths (p = .005), and did not lose their ability to accurately compress between nine and 12 months after training. Oermann, Kardong-Edgren, Odom-Maryon, & Roberts (2014) later studied the use of brief monthly practice of cardiopulmonary resuscitation (CPR) skills among nursing students (n = 606). While there were no statistically significant differences found in compression rates and depth of compressions between the experimental and control groups, there was increased skill retention over a 12-month period. Sutton (2011) had also reported that a six-month post-test retention of skills 2.9

times more likely for providers receiving three brief training sessions in bedside CPR. Vadnais et al., (2012) noted that single day, multiple-task simulation training in critical obstetrical events increased competence retention at four and twelve months post-training.

To promote retention of advanced cardiac life support skills, Wayne and colleagues (2012) provided medical residents (n = 19) with small group teaching, on-line modules, and a log of code status discussions performed by experts in the clinical setting. Compared to the control group who received only clinical training, the intervention group performed significantly better ($p < .001$) on mean scores for a code status discussion and clinical skills examination immediately post-intervention, and one year later.

Forgetting Knowledge and Skills after Training

Historically, in 1913, Ebbinghaus identified a forgetting curve that describes how information can be forgotten after the initial learning, as well as how forgetting can be impacted by the way the information was learned and how frequently the information is reviewed. The Ebbinghaus forgetting curve hypothesizes the decline of retention over time in the absence of intentional attempts to retain the knowledge; this theory continues today.

There are many reports of skill and knowledge forgetting in the training of resuscitation skills. De Regge, Calle, De Paepe, and Monsieurs (2008) reported a decline in skill retention after nine months of CPR training by nurses, providers, and laypersons. Isbye and colleagues (2008) also reported a decline in CPR skills to a pre-instruction level after three months. Furthermore, Smith, Gilcreast, and Pierce (2008) reported that 37% of nurses (*n* = 133) declined in skill retention and were unable to pass a BLS skills test three months post-training, with additional decline noted at six and nine months. Sankar, Vijayakanthi, Sankar, and Dubey (2013) reported declines in both knowledge and skills of pediatric CPR among nurses (*n* = 28) and final semester nursing students (*n* = 46) six weeks post-training. Additionally, Nori, Saghafinia, Motamedi, and Hosseini (2012) reported declines in CPR knowledge and skills at 10 weeks and 2 years post-training among nurses (*n* = 112).

Training retention issues are not limited to the discipline of nursing. Kopacek and colleagues (2010) also reported a decline in learning related to sudden cardiac arrest and automated external defibrillator therapy in pharmacy students four months after training. Shiyovich, Statlender, Abu-Tailakh, Plakht, Shrot, & Kassirer (2015) found that medical personnel's treatment of chemical warfare agent casualties deteriorated significantly before even one year passed since their training.

The importance of training and re-training is evident in these reports of the knowledge and skill sets within a variety of contexts. Absent from the literature, however, are reports of research on the impact of the training of debriefers in nursing simulation education. It is unknown if the type, source, timing, or repetition of training has an impact on a debriefer's understanding and application of an evidence-based debriefing method. While there are numerous descriptions of training in other contexts, there are few reports describing the training of a debriefer in nursing education literature. DML is one evidence-debriefing method that has demonstrated improved clinical reasoning among nursing students is DML (Dreifuerst, 2012). What remains unknown is how debriefers apply this method during debriefing, and if this application is impacted by DML training.

Debriefing for Meaningful Learning©

DML is a structured, evidence-based debriefing method that has been used increasingly by nurse educators in prelicensure nursing education (Dreifuerst & Decker, 2012). While other debriefing methods broadly engage learners in components of reflective thinking and Socratic questioning, DML uniquely facilitates the deepening of thinking processes of learners (Dreifuerst, 2012). By uncovering and guiding learners through a reflective dialogue to reveal thinking, the debriefer empowers the learner to individually expose and analyze the thoughts, feelings, and mental frames underlying their decisions and actions. The development of clinical reasoning is facilitated in learners by actively engaging them in synthesizing, hypothesizing, generalizing, inferring, and questioning (Dreifuerst, 2010). DML is the debriefing method underpinning this study; following is a review of the central teaching-learning concepts embedded within the DML framework, measurement of DML implementation by debriefers, and implications of debriefer training in DML.

Socratic Questioning

As an integral component of DML, Socratic questioning is used to facilitate a conversation in which a debriefer guides a learner to gain understanding of what the learner is thinking by using deliberate questioning (Dreifuerst, 2010).

Systematic questions are posed by the debriefer to deeply assess the truth of the learner's statements, revealing the thought processes and assumptions underlying the learner's statements (Elder & Paul, 1998). The debriefer treats all statements as connecting points to further thoughts, and regards all thoughts as opportunity for development.

The art of Socratic questioning is central to uncovering, correcting, and, developing thinking within the learner (Elder & Paul, 1998). Historically, Socratic questioning was a type of repetitive questioning to prove ignorance and publicly humiliate those who were questioned (Overholser, 1993). Over time, Socratic questioning was refined to promote self-awareness and to develop an open mind (Schmid, 1983) through collaborative exploration (Klein, 1986).

Socratic questioning differs from other types of questioning through the systematic, disciplined, and in-depth style of questioning for the purpose of assessing the plausibility and truth of underlying, thoughts (Elder & Paul, 1998; Overholser, 1993). The goal of Socratic questioning is to guide the learner in evaluating their thinking and resulting actions. The debriefer responds to all answers provided by the learner with a deeper question that provokes the learner to develop his or her thinking more fully. In this manner, all questions are presupposed with prior questions, as all thinking presupposes prior thinking. As learner thinking is exposed through Socratic questioning, both the debriefer and learner gain a more comprehensive view of how the learner is thinking.

Challenging taken-for-granted assumptions. A debriefer can guide learners in challenging taken-for-granted assumptions through Socratic questioning, thus exposing the relationships between thinking and actions (Dreifuerst, 2015). As learners apply new knowledge to learning experiences, Socratic questioning exposes faulty assumptions by uncovering answers, bringing learners to an awareness of their limited knowledge. Through "who, what, where, when, how, and why" questions, the debriefer stimulates learners to self-discover thought processes and connect their thinking to actions performed during the experience. Recognizing that each thought does not exist in isolation, but is nested within a network of connected thoughts, the debriefer guides the learner in discovering and examining these connections. Often, learners' existing mental frameworks, and resulting connections of thoughts, are founded on faulty assumptions. Through engaging learners in Socratic dialogue, a debriefer exposes faulty assumptions and guides learners to reframe their thinking.

Reflection

To reflect is to think purposefully (Dewey, 1910), and to intentionally observe and notice one's thinking and actions (Pesut & Herman, 1999). Dewey (1910) first introduced the concept of reflection as purposeful thought, later describing the importance of reflection to critical thinking as "the turning over of a subject in the mind and giving it serious and consecutive consideration" (1933, p. 23). Mezirow (1981) deepened this description by articulating the reflective process as "becoming aware of our awareness" (p. 13) and described stages of a reflective process to make meaning from experiences.

Schön (1983) extended these descriptions of reflection as a result of his attempt to understand and explain the thinking of professional, expert practitioners. Schön described a practitioner's capability of "thinking on your feet", suggesting the aptitude to simultaneously think about an act while doing it. He argued that the technical rationality model traditionally prominent in professional training focused on knowledge rather than "thinking in action". Schön's (1983) articulation of reflective practice described reflection as the process professional practitioners engage in when facing complex situations which cannot be solved solely by technical rational methods, demanding the need for reflection in practice.

Reflection was defined by Boyd and Fales (1983) as an internal process of examining an experience, for the purpose of clarifying meaning. Kolb (1984) included the stage of reflective observation in the experiential learning cycle, as a means for deepening a learning experience by observing and critically examining the why and how of the experience. Boud, Keogh, and Walker (1985) described reflection after learning experiences as the opportunity to explore experiences to gain new understanding.

While reflective thinking may not occur innately, it can be developed through instruction and modeling over time (Decker et al., 2013). Facilitating reflection among learners, or guided reflection, promotes the development of clinical reasoning, a necessary component of learning to think like a nurse (Dreifuerst, 2009), and teaches students to frame clinical experiences correctly (Dreifuerst, 2010; Murphy, 2004). Reflection is engaging in thinking about the underlying causes of one's thoughts, feelings, and behaviors related to an experience. The insight and understanding gained through reflective thinking can facilitate purposeful cognitive, affective, and behavioral change. Reflection is paramount to metacognition, a higher order cognitive process described as thinking about one's thinking (Kuiper & Pesut, 2004).

While many teaching-learning methods have been used to teach reflection within nursing education, guided reflection is helpful for novice practitioners who may be lacking in the skills needed to reflect deeply (Kuiper & Pesut, 2004). A learner will be more likely to engage in reflection when they understand and embrace the impact of reflection on individual and professional development (Scanlan & Chernomas, 1997). Atkins and Murphy (2008) identified learning as an outcome of reflection, while others identified an additional outcome of reflection as a purposeful change in thinking and behaviors (Boyd & Fales, 1983; Boud, Keogh, & Walker, 1985).

Schön (1987) conceptualized three stages of reflection that have been found useful in nursing education (Atkins & Murphy; Kuiper & Pesut, 2004; Scanlan & Chernomas, 1997), and are specifically stated as the phases of debriefing in INACSL Standard VI (Decker et al., 2013): awareness, analysis, and summary. The first stage, awareness, is described as a response to thoughts or feelings arising from an unexpected or surprising situation. This beginning awareness of thoughts and feelings is important in raising curiosity, and fostering an inquisitive spirit. During the second stage of reflection, the situation is critically analyzed, then a new perspective of the situation is developed through the analysis of what happened and the application of new insight and information to the learning experience during the final stage of reflection. Three types of reflection are reviewed that are supported by these phases, as well as further expounding each phase.

Reflection-in-action. Reflection in and on action are significant components of reflective practice. Schön (1983) differentiated between reflection occurring during practice, or reflection-in-action, and reflection occurring while thinking back on practice, or reflection-on-action. Reflection-in-action describes a thinking that facilitates reshaping of judgments, decisions, and experience during practice while reflection-on-action achieves this retrospectively (Dreifuerst, 2015). Schön's description of these two types of reflection were based on descriptions of the types of reflection practitioners engage in. However, practitioners are often not consciously aware of the knowledge used while reflecting in action, particularly novices and students. Therefore, it may be difficult to articulate the knowledge and the thought processes they are using in-action. Kuiper and Pesut (2004) reiterated the value of structured reflection to novice practitioners who may be lacking in the skills necessary for analyzing practice in the moment; retrospective reflection may be used to teach novices to recall all aspects of active thinking in the moment to learn how to reflect-in-action.

Engaging learners in reflection-in-action guides them in examining thinking and decision-making processes used in the moment during a learning experience (Dreifuerst, 2015). By its very nature, this is difficult to teach and to learn because it occurs as-lived. Since debriefing occurs after the experience, the debriefer guides learners in an exercise of coming back into the moment even though in reality it is retrospective reflection-in-action. Reflection-in-action during debriefing is different from reflection-on-action as the debriefer guides the learner to reflect on being in the moment while the nursing actions and decisions were occurring. Guided reflection not only deepens the learner's understanding of the learning experience, but also promotes the learner's awareness of reflection-in-action as it occurs in future experiences.

Likewise, guiding learners through reflection-in-action during debriefing allows for a thorough examination of the thoughts that occur before and during the action (Dreifuerst, 2015). This act of making thinking in the moment evident is important for two reasons. First, learners develop an awareness of thinking in the moment that they can carry into nursing practice. Second, the ability to reflect-in-action guides learners as they develop into practitioners in the process of challenging their own taken-for-granted assumptions in real-time which may promote a safer approach to patient care.

Reflection-on-action. Reflection-on-action is reflection that occurs after an action (Dreifuerst, 2015). This type of reflection is a time of critical reexamination of the thinking processes, feelings, and factors impacting decisions and actions before, during and after an experience (Schön, 1983; Tanner, 2006). Reflection-on-action is a review of an experience, and often involves a sense of inquisitive wondering about possible outcomes that may have resulted from other decisions or actions that could have been made during the experience. It is also a time when things forgotten in the moment may become apparent in the conscious thoughts of the practitioner as the debriefer guides the participant through the process of uncovering the thinking behind the actions that occurred.

While this type of reflection may happen innately, the debriefer guides the learner through a verbal recounting of reflection-on-action to analyze thoughts and actions collaboratively. As the learner verbalizes thought processes, past experiences, and other contributing factors to the decisions and actions made during the learning experience, the debriefer can uncover correct and incorrect thinking, facilitate deeper thinking in the learner, guide the learner to self-correct and reframe where necessary, or affirm correct thinking and actions. Guided reflection-on-action develops this reflective skill set in learners which will promote reflective activity in nursing practice. Reflective thinking promotes a safer approach to nursing care since the ability to reflect-on-action during nursing practice develops over time and is a hallmark of developing expertise (Benner, 1984; Kuiper & Pesut, 2004).

Reflection-beyond-action. Dreifuerst (2009) extended Schön's (1983) definition of reflection to include reflection-beyond-action. The ability to anticipate is closely related to reflection. Reflection-beyond-action involves examining what has occurred while anticipating how this learning may be applied the next time a similar clinical situation is encountered. Through a guided reflection-beyond-action, learners can begin to connect nursing care provided in specific clinical contexts to a broader scope of patients with guidance from a debriefer, who is also a subject matter expert. Reflection-beyond-action is an important thinking skill to teach to novices, particularly because of the limited number of clinical situations presented within a nursing program. Guiding learners through reflection-beyond-action extends thinking beyond the current situation, which is an important skill in applying current learning to future similar situations and teaching anticipation, another hallmark of developing expertise (Benner, 1984; Kuiper & Pesut, 2004).

Six E's of DML

The six E's of DML were adapted from the Biological Sciences Curriculum Study (BSCS) E5 Instructional Model, whose origin extends back to math and science instructional models used since the early 1900s (Bybee et al., 1989; 2006; Dreifuerst, 2010). The BSCS 5E Instructional Model consists of the phases engagement, exploration, explanation, elaboration, and evaluation. Through this instructional model, learning is viewed as dynamic and interactive; learners collaborate with teachers to construct new knowledge, building on prior knowledge and experience to construct meaning of challenging concepts, and provide opportunities for reconstruction when necessary. The five phases of the BSCS 5E Instructional Model were extended by Dreifuerst (2010) with an additional sixth phase of extend to formulate a model that fosters anticipatory thinking.

Engage. During the initial phase of engage, the DML debriefer facilitates learner engagement by gathering the learners to begin debriefing after the learning experience (Dreifuerst, 2010). Learners are invited to silently journal initial thoughts and release their feelings from the learning experience. During this phase, the debriefer leads the learners through framing the patient situation and identifying a focused key problem.

Although the debriefer engages students in the beginning stages of debriefing, engagement is not exclusive to the first phase of the debriefing session. Rather, all of the six E's are iteratively present throughout a DML debriefing. In this manner, engagement is enacted by the debriefer through maintaining a listening posture, redirecting conversation when necessary to keep all learners engaged, and abstaining from a lecture format throughout the debriefing session. The debriefer facilitates an intellectually engaging environment through Socratic questioning, generating further exploration of thought, moving learning forward throughout the debriefing session.

Explore. In the explore phase, the debriefer guides learners in an exploration of the clinical decisions that were made, as well as identifying contributing factors in making these decisions (Dreifuerst, 2010). Through engagement in Socratic dialogue, the debriefer guides learners to reflect-in-action and reflect-on-action in order to uncover the thinking that elicited an action, or a decision to not act. The debriefer asks questions to guide learners in making connections between thinking, decisions, and actions. During the explore phase, the debriefer challenges the learner's taken-for-granted assumptions to further explore the reasoning behind the nursing actions apparent or not apparent in the learning experience. The debriefer also guides the learner in identifying what could have been done differently in the learning experience.

Explain. During the explain phase, the debriefer guides learners through explaining patient assessments, decisions made, and decision-making processes in the learning experience, as well as patient responses to decisions made (Dreifuerst, 2010). The debriefer guides learners in making connections between thinking and actions, and to think about connecting the care of the patient in the current learning experience to a broader scope of these types of patients. The debriefer may also need to explain clinical knowledge when the learner's knowledge or assumptions are incorrect.

Elaborate. Through elaboration of thoughts, behaviors, and isolated components of the learning experience, the debriefer facilitates a conversation that expands learner thinking and actions (Dreifuerst, 2010). During this phase, the debriefer guides learners to identify critical points of the learning experience, as well as verbalizing what went right and what went wrong. In addition, the debriefer facilitates conversation that guides learners to further elaborate on thoughts and feelings underlying all thoughts and actions that occurred during the learning experience. During this phase, the debriefer guides the learners to elaborate on their assumptions, while expanding the relationships between thinking and actions through reflection in and on action.

Evaluate. Through continued guidance by the debriefer in reflecting on action, learners and the debriefer collaboratively evaluate what did not go well in the learning experience (Dreifuerst, 2010). Through guided reflection, learners can evaluate their knowledge, decisions, actions, and resulting patient outcomes. This phase is essential in re-framing the learning experience so that the experience is cognitively locked into memory with the appropriate decisions and actions. Through guiding learners in evaluation of all aspects of their thoughts and actions in the learning experience, the debriefer is able to collaboratively confront misconceptions together with the learner by revealing their thinking and assumptions in order to restructure their mental frames.

Extend. In the final phase, the debriefer encourages learners to extend their thinking beyond the isolated learning experience to a future possible encounter. This is achieved through guided anticipation and reflection-beyond-action (Dreifuerst, 2010). Using 'what if' questions, debriefers guide learners in thinking beyond the confines of one isolated clinical encounter to see how many concepts involved in the care of this patient may also apply to the care of other patients in both similar and different circumstances. As learner thinking is exposed, the debriefer can help the learner continually push thinking forward to imagine future unexpected and unanticipated situations. In addition, debriefers equip learners to take charge of developing their own understanding with guidance of an expert in clinical knowledge.

Engaging learners in guided reflection through Socratic dialogue is an essential debriefing skill for applying DML. In tandem with reflective thinking, the thinking constructs embedded in each of the DML phases collectively promote the development of clinical reasoning, or thinking like a nurse (Dreifuerst, 2012).

Thinking Like a Nurse

Learning to think like a nurse is significant to the development of clinical reasoning in nursing students (Tanner, 2006). The need to teach a student to think like a nurse was a crucial finding of the Carnegie Foundation for the Advancement of Teaching's national nursing education study (Benner, Sutphen, Leonard, & Day, 2010). This finding led to a call for a shift from emphasizing the traditionally upheld thinking style of critical thinking to an emphasis on clinical reasoning, to include contextual, higher order reasoning methods.

The literature reflects disciplinary consensus that clinical reasoning is comprised of higher order thinking processes a nurse engages in to make clinical decisions (Banning, 2008; Tanner, 2006; Kautz, Kuiper, Pesut, Knight-Brown, & Daneker, 2005; Kuiper, Pesut, & Kautz, 2009; Lapkin, Levett-Jones, Bellchambers, & Fernandez, 2010; Levett-Jones, et al., 2010, Murphy, 2004; Simmons, 2010). Clinical reasoning has been distinguished from lower levels of thinking as an essential component of competent nursing practice, and comprises the highest order level of complex thinking a nurse engages in (Dreifuerst, 2012; Lapkin et al., 2010; Simmons, Lanuza, Fonteyn, Hicks, & Holm, 2003).

In any clinical situation, a nurse engages in multiple non-linear thinking routines simultaneously. Thinking skills are the building blocks of thinking routines (Ritchhart, Church, & Morrison, 2011). In isolation, any of these thinking skills alone does not constitute clinical reasoning or thinking like a nurse. Thinking routines in a clinical context are the building blocks of clinical reasoning. As each thinking skill is learned, thinking routines are fostered through intentional instruction and practice; ultimately, clinical reasoning develops increasingly as nursing students apply these thinking routines to a clinical context.

Use of DML in Nursing Education

The use of DML in debriefing has been a focus of research in nursing education to promote clinical reasoning, or the ability to think like a nurse. Dreifuerst (2012) investigated the relationship between the use of the DML method and the development of clinical reasoning skills in prelicensure nursing students (n = 238) and found a statistically significant difference ($p < .05$) between pretest and posttest Health Science Reasoning Test (HSRT) scores. In addition, students perceived a difference in the quality of debriefing when the DML method was used, compared to the absence of the DML method by the debriefer. Furthermore, higher student perception of the quality of debriefing was associated with greater positive changes in HSRT posttest scores.

Mariani and colleagues (2013) compared the clinical judgment development of baccalaureate nursing students (n = 86) who participated in debriefing sessions with the DML method, and of students who participated in unstructured debriefing methods. The Lasater Clinical Judgment Rubric (LCJR) was used to measure the influence of structured debriefing with DML on clinical judgment. Clinical judgment was defined as a complex skill developed through decision making, critical thinking, clinical decision making, and clinical reasoning. While there was not a statistically significant difference between the groups on the LCJR scores, focus group interviews of the participants reveal participants who were debriefed with DML perceived this type of debriefing to foster a learner-centered, holistic approach that helped them map and analyze meaningful concepts that would support their knowledge and thinking for future situations.

The DML method was the chosen debriefing method used in the longitudinal, multisite NCSBN NSS study that tested if substituting clinical hours with 25% and 50% simulation with prelicensure nursing students (*n* = 847) impacted knowledge, clinical competency, critical thinking, and readiness for practice (Jeffries et al., 2015; Hayden et al., 2014). Hayden and colleagues (2014) reported that up to 50 percent simulation hours could be substituted effectively for traditional clinical experiences throughout the nursing curriculum when conditions were similar.

Forneris and colleagues (2015) sought to replicate the findings of Dreifuerst's (2012) research in a multi-site study testing the impact of DML on clinical reasoning in prelicensure baccalaureate nursing students (*n* = 153). The HSRT was used during the first week of class, and three weeks after the intervention to measure changes in clinical reasoning. The intervention included a simulation experience from the NLN's Advancing Care Excellence for Seniors (ACES) scenarios, followed by a DML debriefing facilitated by debriefers trained in the method. The change in the mean HSRT score for students in the intervention group were statistically significant at the .05 level (*p* = .03); the change in the mean HSRT score between the intervention and control groups was statistically significant (*p* = .09) at the .10 level. Nursing students demonstrated an improvement in clinical reasoning when debriefed when DML, compared to students who were debriefed in a customary debriefing.

Testing of DML in nursing education has demonstrated improved development of clinical reasoning. However, there is a scarcity of literature describing a debriefer's understanding of the central concepts of the method, and subsequent consistent application of the method as it was originally designed and tested. Also absent from the literature, is a description of the impact of training in DML on how debriefers apply the method in debriefing.

Evaluation of the Debriefer

Debriefing is one increasingly common teaching and learning process during which teaching can be assessed and evaluated (Dreifuerst, 2015). Traditionally, evaluation of debriefing has centered on examining components of debriefing facilitation such as length of time, environment, faculty experience, and learning objectives (Waznonis, 2015). While these constructs are essential to the debriefing experience, the skill of the debriefer applying a debriefing method has received little evaluation.

In recent debriefing surveys, Fey (2014) reported that 19% (*n* = 92/484) of respondents reported debriefing was assessed; 81.5% (*n* = 75/92) of respondents reported that this was done through observation and feedback without the use of an instrument. The most frequently reported debriefing evaluation instrument used (*n* = 23/92 or 30.6%) was the DASH-RV© (Center for Medical Simulation, 2014). Waznonis (2015) reported similar findings. Of respondents who reported evaluation of debriefing sessions (*n* = 65/205), approximately half reported evaluation with an instrument and half received evaluative feedback by another unnamed method (Waznonis, 2015). Of the respondents reporting evaluation with an instrument (*n* = 24/42), respondents described the use of an evaluation instrument, or a self-developed rubric to evaluate debriefing effectiveness. The most frequently reported debriefing evaluation instrument used was the DASH-RV© (Center for Medical Simulation, 2014), reportedly 40% (*n* = 16/40). Other reported evaluation methods included student feedback (89%), self-reflection (74%), instructor feedback (69%), observer feedback (37%), and other student assignments and activities (29%).

Debriefing Instruments

In a literature review of methods and evaluation of simulation debriefing, Waznonis (2014) reported seven evaluation measures of debriefing. Only six of the reports identified an instrument as the means for evaluating the effectiveness of simulation debriefing. Each of the six instruments were intended to broadly evaluate the debriefing experience; three of the reported instruments were designed for peer review, three instruments were designed to be used by debriefing participants, and one instrument was designed for self-evaluation (Waznonis, 2014). Three instruments were reported to be designed for use by peer reviewers for rating dimensions of the debriefing experience: a 25-item instrument for rating debriefing effectiveness (Guraraja, Yang, Paige, Shauvin, 2008), the Objective Structured Assessment of Debriefing (OSAD; Arora et al., 2012), and the DASH© (Simon et al., 2010a).

The Guraraja et al. (2008) instrument was designed with 25 items representative of effective debriefing environments, not specific to setting, participant, or content. This instrument was used to rate debriefing effectiveness after point of care simulations with video-recorded debriefing discussions. The instrument was tested in conjunction with the System for Teamwork Effectiveness and Patient Safety (STEPS) training program in a mobile mock operating room model. Reporting of testing of the instrument included mean scores for each of the items of the 25 Likert-type response scale. There is no reporting of testing of debriefing with a specific debriefing method.

The DASH-RV© was developed with the purpose of evaluating debriefing quality across disciplines and in a variety of simulation learning environments, and is not specific to any debriefing method (Simon, et al., 2010a). The DASH-RV© is a behaviorally anchored rating scale (BARS) comprised of 6 elements; each element is a concept describing an area of the debriefing experience, elaborated by dimensions of observable debriefing behaviors. The items of the DASH-RV© were developed based on behaviors accepted as best practices for effective debriefing, as well as behaviors indicated by theory as facilitating learning and change (Brett-Fleegler et al., 2012). Content experts reviewed items for validity, prior to pilot testing by 151 international health care educators. Interrater reliability was assessed, and variance component analysis was used to calculate interclass correlation coefficients (ICC); total scale ICC was calculated as 0.74. Internal consistency of the scale was determined by evaluation of Cronbach's alpha, calculated to be 0.89.

There are two other versions of the DASH: the student version (DASH-SV©) to be used by students for rating debriefers (Simon, Raemer, & Rudolph, 2010b), and the instructor version (DASH-IV) to be used for debriefer self-assessment (Simon, Raemer, & Rudolph, 2010c). Brett-Fleegler and colleagues (2012) noted limitations of the DASH-RV© included the lack of psychometrics outside of this setting, and that validity testing is needed for use in different simulation settings and with different raters than the participants of the initial psychometric analysis. In addition, further investigation was noted as necessary to determine how well the DASH-RV© can be used to assess different debriefing methods.

The DASH-SV© Long Form was designed to be used by participants to rate the debriefer on the six DASH© elements and the 23 behaviors within each element. The DASH-SV© Short Form was designed for participants to rate the debriefer on only the six DASH© elements (Simon, Raemer, & Rudolph, 2010d). One additional reported measure of self-evaluation is the DASH-IV© (Simon et al., 2010c). The DASH-IV© Long Form was intended for debriefers to rate the six DASH© elements and the 23 behaviors within the six elements. The DASH-IV© Short Form provides a self-evaluation of six DASH© elements (Simon, Raemer, & Rudolph, 2010e).

The OSAD tool was developed to assess debriefing quality focusing on key elements of an effective debriefing (Arora et al., 2012). The items of this tool consist of descriptive anchors to guide rating of the scale. Evidence of content validity (Content Validity Index = 0.87) and concurrent validity were reported with testing of physicians debriefing surgical residents. While testing of the OSAD generated reliable data for assessing simulation debriefing, testing of these instruments was limited to a sample of surgical residents. As with the 25-item instrument from Gururaja and colleagues (2008), there is no report of testing with a specific debriefing method (Arora et al., 2012).

The Debriefing Evaluation Scale (DES) was designed by Reed (2012) to be used by debriefing participants to describe the nursing student's experience during debriefing. Two scales were developed within the DES; the student experience during debriefing, and the student's perception of the importance of the experience. A 20-item Likert scale format was used to rate perceptions of the debriefing experience for the scale. Factor analysis was completed and internal consistency was determined by Cronbach's alpha, and was calculated to be .91 for the experience scale. No psychometric testing was performed on the importance scale. The DES was not designed to evaluate a specific debriefing method.

The Debriefing for Meaningful Learning Student Questionnaire (DMLSQ) was developed by Dreifuerst (2012) for the purpose of gathering feedback from participants during the development and testing of DML, in order to inform the investigator of the perceptions of the participants when debriefed with the DML method. The DMLSQ consisted of four questions focused on assessing the usefulness of the DML worksheets, the participants' perception of their learning from debriefing, and perception of the time allotted for debriefing. It has not been used since the initial development of the debriefing method.

Although there are multiple endorsements by INACSL, the NCSBN, and the NLN regarding the value of adhering to an evidence-based, structured debriefing method, there are no guidelines regarding the use of an instrument to measure how well a debriefer implements a specific debriefing method. Furthermore, many of the instruments reportedly used to evaluate the quality of the debriefing experience rely on subjective rating of behaviors, or components of the debriefing session. An objective measure is needed to provide specific data on each debriefer's behaviors during debriefing. This is a critical initial step in order to evaluate how a debriefer engages learners through a debriefing method; this data could enrich the preparation and training of debriefers in evidence-based debriefing methods. and could explain student outcomes from debriefing. If a particular debriefing method is used to attain certain learning outcomes following an experience, then it is vital to have a method for evaluation of adherence to, and effectiveness of, each debriefer's implementation of that method.

Despite the adoption of DML by faculty in many schools of nursing, it is not known how this method is used in debriefing with prelicensure baccalaureate nursing students. Therefore, a behaviorally anchored rating scale (BARS) based on the iterative process of DML was developed by the investigator to assess a debriefer's implementation of this debriefing method (Bradley & Dreifuerst, 2016). The Debriefing for Meaningful Learning Evaluation Scale (DMLES) was developed to measure use of the elements and process of DML by debriefers during simulation debriefing. The DMLES is a 31-item scale behaviorally anchored in the iterative process of DML debriefing (Bradley & Dreifuerst, 2016). The scale was developed to measure how debriefers implement DML in simulation debriefing with prelicensure nursing students. The observable behaviors of the six E's of DML (engage, explore, explain, elaborate, evaluate and extend) were structured as 31 distinct items of the DMLES.

Face validity was assessed by the developer of DML and two known experts in the method (Bradley & Dreifuerst, 2016). All three agreed the scale was representative of the DML method. Content validity was quantified through the Content Validity Index (CVI) for each item (I-CVI), which was then summed and divided by the total number of items, yielding the Scale CVI (S-CVI) of .92. An acceptable level is an index of .80 or greater (Lynn, 1986). In initial testing of the scale, the DMLES demonstrated high reliability with a Cronbach's alpha coefficient of .88 for the total scale. Interrater reliability was determined through intraclass correlation coefficients (ICC) which was .86 for the total scale ($p < 0.01$).

Summary

Debriefing is the most significant component of a simulation experience, (Shinnick et al., 2011), yet little is known regarding the training of debriefers for this critical learning time. DML is one evidence-based debriefing method that has demonstrated improvement in learner's ability to think like a nurse (Dreifuerst, 2012; Forneris et al., 2015; Mariani et al., 2013). Because of this, DML has been widely adopted for use in curriculum throughout nursing education. However, little is known regarding how DML is being used in debriefing.

Within simulation debriefing literature, seven instruments for evaluating debriefing have been reported; each were intended to broadly evaluate the debriefing experience (Waznonis, 2014). Despite the existence of these instruments for evaluating debriefing, only 20 to 30 percent of debriefers have reported regular competency assessment of debriefing practice (Fey, 2014; Waznonis, 2015). While these instruments are available for measuring debriefing quality or the debriefing environment, there are no reports of instruments measuring the specific skill of a debriefer to implement a structured debriefing method.

Because of the lack of debriefing evaluation, the scarcity of instruments used to measure debriefing skills, and the widespread adoption of DML within nursing education, the DMLES was developed to measure a debriefer's implementation of DML (Bradley & Dreifuerst, 2016). The pilot testing of the DMLES provided internal consistency, interrater reliability, face validity, and content validity data to continue the use of this instrument.

Furthermore, literature supports the use of the DMLES in measuring the observable behaviors of DML, an evidence-based debriefing method, to provide a closer examination of the effectiveness of this debriefing method (Bradley & Dreifuerst, 2016). Since there are many unanswered questions regarding how DML debriefing is implemented, this study will describe the impact of various types and sources of training on application of DML, and add to the body of literature of simulation debriefing.

Chapter III Methodology

The purpose of this study was to describe the impact of debriefing training on how consistently debriefers understand and apply the central concepts of Debriefing for Meaningful Learning© (DML), when grouped according to the type of training received. Participants completed an electronic survey describing their typical simulation debriefing with prelicensure baccalaureate nursing students. This chapter summarizes the methodology implemented in this study including the selection of participants, instrumentation, data collection, and data analysis.

Selection of Participants

Debriefers who facilitate simulation debriefing with prelicensure baccalaureate nursing students were recruited to participate in this study. Participants were included in the study if they were nurse educators who were currently teaching in a baccalaureate nursing program, and reported having received training in debriefing. A priori, the desired sample size was determined by a power analysis using G*Power© (Faul, Erdfelder, Lang, & Buchner, 2007). Based on a power analysis with $p < 0.05$, a power of 0.80, and a moderate effect size, a total sample size of 127 was determined to be needed.

Sampling Procedures

Upon approval from the Institutional Review Board (Appendix A), a Study Information Sheet (SIS; Appendix B) with a Uniform Resource Locator (URL) was posted on the LinkedIn page for the International Nursing Association for Clinical Simulation and Learning (INACSL), in order to attract nurse educators that engage in simulation debriefing as study participants. Members of INACSL who were known debriefers in prelicensure baccalaureate nursing programs were also emailed an invitation with the URL to the SIS and survey to broaden recruitment, as not all members regularly access the LinkedIn group.

After clicking the URL, potential participants were directed to the SIS which described the study and explained that participation was anonymous, voluntary, and confidential. Agreement to participate then prompted the respondents to complete demographic questions, followed by a 57 item survey (Appendix C), previously uploaded into Survey Monkey© by the investigator.

Sample

Of the 308 respondents who accessed the survey link, 287 met the inclusion criteria of teaching prelicensure baccalaureate nursing students. Of the 287 respondents who met inclusion criteria, 283 agreed to participate, one exited the survey, and three declined to participate. Three participants exited the survey prior to completing demographic information, resulting in 276 participants. Table 1 summarizes the demographics of the debriefers who met the inclusion criteria of the study.

Table 1

Debriefer Demographics

Demographics	*N*	Frequency
Type of teaching institution[a]		
University/college	274	(99%)
Hospital-based nursing program	2	(1%)
Highest academic degree completed[a]		
Bachelor's degree	26	(9%)
Master's degree	165	(60%)
Doctoral degree	85	(31%)
Years teaching traditional BSN students[b]		
Less than 1 year	16	(6%)
1-5 years	73	(27%)
6-10 years	84	(31%)
11-15 years	49	(18%)
16-20 years	27	(10%)
21 or more years	26	(10%)
Years using debriefing in simulation[b]		
Less than 1	17	(6%)
1-2	38	(14%)
3-4	58	(21%)
5-6	64	(23%)
7-8	47	(17%)
9-10	27	(10%)
10 or more	24	(9%)

[a]$n = 276$
[b]$n = 275$

Instrumentation

This study sought to describe the impact of training on how debriefers implement DML in simulation debriefing with prelicensure baccalaureate nursing students. The initial step was to quantify the behaviors consistent with a DML debriefing. Without first understanding how DML is implemented in debriefing, there is limited ability to measure interactions with different sources of training. Because there was no existing tool that measures how a debriefer implements DML, the Debriefing for Meaningful Learning Evaluation Scale (DMLES; Appendix D) was developed and pilot tested by the investigator to be used as an observational rating scale of a debriefer's implementation of DML with baccalaureate nursing students (Bradley & Dreifuerst, 2016).

The DLMES was designed to be used by raters who observed debriefers and rated their use of DML. Unfortunately, recruitment of debriefers who would agree to participate in a study in which their debriefing was observed and rated proved too challenging to be feasible for a research study. Therefore, a self-report survey, the Debriefing for Meaningful Learning Inventory (DMLI), based on the items previously developed for the DMLES, was designed for this study. The psychometric testing of the DMLES, as well as the development and testing of the DMLI, will be described in the subsequent sections.

Debriefing for Meaningful Learning Evaluation Scale (DMLES)

Measurement is a scientific activity that quantifies observations about people, processes, events, or attributes (DeVellis, 2012; Thorndike & Thorndike-Christ, 2010). Through identifying the presence of observable debriefing behaviors aligned with the DML debriefing method (Dreifuerst, 2012), the Debriefing for Meaningful Learning Evaluation Scale (DMLES) was developed to measure the ability of a debriefer to implement DML (Bradley & Dreifuerst, 2016).

Scale framework. Item Response Theory (IRT) has been used extensively in educational assessment such as the GRE and the Stanford-Binet 5 Intelligence scales (DeChamplain, 2010; Warne, McKyer, & Smith, 2012), licensure and certification frameworks (DeChamplain, 2010), and health research (Warne, McKyer, & Smith, 2012). In healthcare, IRT scales have been used with instruments measuring behavioral constructs including smoking behavior, nursing self-efficacy, and anxiety. An advantage of IRT is the provision of greater detail on the precision of a measure (DeChamplain, 2010; Nguyen, Han, & Chan, 2014).

IRT guided the development of the DMLES due to its focus on behavioral, item-level detail. The purpose of IRT models is to explain observed item performance as a measure of a latent trait (DeChamplain, 2010). IRT comprises a group of non-linear, mathematical models that use a pattern of responses to a group of items to make predictions about latent variables that the items measure (Warne, McKyer, & Smith, 2012). IRT models describe the relationship between an individual's ability to respond to scaled items and focuses on item-level information (Raju, Su, & Patrician, 2014). IRT is useful in identifying best items to use on the basis of the purpose of a measure, identifying where an item performs best, and identifying measure equivalence across subgroups (Nguyen, Han, Kim, & Chan, 2014). Key assumptions of IRT include unidimensionality of the measured traits, or the assumption that a set of items measure one common trait. Additional IRT assumptions include local independence (every item is statistically independent of responses to other items), monotonicity (the probability that selecting an item will increase as the trait level increases), and item invariance (estimated item parameters are constant across populations).

Scale development. Generating a set of similar or related items does not constitute a scale (DeVellis, 2012). A necessary component of a scale is to determine the characteristics of the latent variables underlying an item set. An observational scale was chosen for the DMLES for the purpose of collecting direct information of a debriefer's behaviors during direct observation of a debriefer implementing DML (Dowdy, Twyford, & Sharkey, 2013). DeVellis (2012) outlined specific guidelines for use in developing measurement scales: 1) define the construct to be measured, 2) generate an item pool, 3) determine measurement format, 4) expert review of the item pool, 5) inclusion of validation items, 6) administer items to a sample, 7) evaluate the items, 8) optimize scale length. Following is a description of the process of the development and testing of the DMLES.

***Define the construct**.* The construct to be measured with this scale was a debriefer's implementation of the overall process and each step of DML. The measurement of abstract latent variables is challenging since they cannot be directly observed or tangibly identified (DeVellis, 2012). Because this scale development was guided by the theoretical definitions of thinking strategies embedded within the DML framework, the abstract concepts of DML were first identified, then further clarified through literature review and finally validated by the DML author. As an established debriefing method, the method of DML renders identifiable steps of a structured debriefing.

Item generation. The items of the DMLES were generated from the iterative steps within each of the six E's of DML. While this generated a larger item pool than would be pragmatic for use in practice, the large pool of items was evaluated for eventual inclusion in the final scale to ensure consideration and testing of all possible items (DeVellis, 2012). Generating items during those initial stages required defining and describing each element as underlying central concepts of the DML framework, and thoroughly describing each element with observable examples of behaviors in each description. Careful consideration of the observable behaviors of each element was necessary to construct distinct items, allowing for independent rating of each item without overlap (Brett-Fleegler et al., 2012). Each item was chosen to reflect the purpose of the scale succinctly, without redundancy or ambiguity (DeVellis, 2012).

Determine measurement format. Teaching is a complex process impacted by multiple contextual factors, making the measurement of any aspect of the teaching process challenging to quantify (Conigliaro & Stratton, 2010). One approach to measuring the ability to teach is to identify observable teaching behaviors (Conigliaro & Stratton, 2010; Dowdy, Twyford, & Sharkey, 2013). A behaviorally anchored rating scale (BARS) is an evidence-based measurement scale that measures performance, particularly related to competence, through combining an objective scale with numeric levels to quantify behaviors (Anitei &

Chraif, 2012; Brett-Fleegler et al., 2012). The behavioral anchors of this instrument are the described iterative steps of the DML method.

The DMLS was developed to score observable behaviors with binary options for scoring: (1) *present* and (2) *not present*. While initial consideration was given to a Likert scale format, the purpose in development was to ascertain whether teaching ability could indeed be quantified. Although a Likert scale format would provide description of the varying degrees of observable behaviors (DeVellis, 2012), the purpose of this scale was to identify the presence or absence of each element of the process of the DML method during a typical debriefing based on DML.

Expert item review. The items of the scale were initially reviewed by the author of DML and two known experts in the method to ensure that each iterative step of the embedded thinking strategies was captured and portrayed within each item of the six elements of the scale. Each dimension of the six elements of DML were also examined for clarity, order, and observable descriptions. Every item was judiciously regarded to minimize overlap between the items. On initial review, 27 items were identified for scale inclusion. After item review with the DML author, 7 items were added to increase clarity and precision of scoring, yielding a total of 34 items for pilot testing. Prior to testing of the scale, face validity was assessed by the developer of the DML method, with subsequent item and total scale revision, yielding the 33-item scale. After testing of the scale, additional validity measures were sought to further assess validity of the constructs represented by the DML method. Face validity was assessed through a questionnaire completed by two additional experts of the DML method. All experts agreed the scale was representative of the DML method. Recommendations were made regarding wording, ordering, and addition of items.

Consider inclusion of validation items. Although distinct concepts, the concepts of reliability and agreement are often used interchangeably (Kottner et al., 2011). Agreement refers to the degree to which scores are identical, while reliability is the ability of a measurement to differentiate between subjects, or the ability of an instrument to measure consistently (Devellis, 2003). Lucas and colleagues (2009) described three types of reliability: diagnostic reliability (agreement between two or more observations of an entity), interrater reliability (agreement between two or more raters observing the same entity), and inter-rater reliability (agreement of two or more observations of the same entity by a single rater). Diagnostic and interrater reliability were important in the development of this instrument in the measurement of the abstract elements of a debriefing method.

Michell (2000) described the need for reliable access to the phenomena under investigation, highlighting the unique ability to give quantifiable value to an abstract concept by assigning numbers to represent the properties of abstract phenomena. Enhancing construct validity within this instrument was a key component to maintaining evidence of the validity of the scale (DeVellis, 2012). Establishing a theoretical relationship between variables and correlating empirical relationships with the predicted patterns through the identified observable behaviors of the DML method provided evidence of how well the items of the scale measured what is intended to be measured.

Administer items to a sample. The DMLES was pilot tested with a sample of three debriefers teaching within a Midwestern prelicensure nursing program who were purposively solicited to submit debriefings for review based on the following criteria: each received training in DML by the developer of DML in conjunction with the National Council of State Boards of Nursing (NCSBN) National Simulation Study, and each facilitated debriefing with prelicensure nursing students During regularly scheduled simulation learning experiences, students and debriefers participated in a 20 minute simulation, followed by a 30 minute debriefing per the customary process within the program. Each debriefing was recorded; each of the three faculty submitted five recorded debriefings to be scored, for a total of 15 recorded debriefing sessions.

Three expert debriefers with extensive DML experience were asked to individually and privately score 15 recorded debriefings with the DMLES after receiving training in use of the scale. Training included review of all components of DML, and instruction regarding the observational approach used to collect direct information about a debriefer's teaching behaviors (Dowdy, Twyford, & Sharkey, 2013).

Evaluate the items. After administration of the 33 items with the representative sample, the next step was evaluation of the performance of each item to identify which items should be included in the final scale (DeVellis, 2012). Because the ultimate goal of each individual item was a high correlation with the true score of the latent variable, and because the true score could not be directly assessed or computed to correlate with items directly, inferences were made based on the measurement model. Highly inter-correlated items within the scale demonstrate higher reliability; the more highly correlated the items were, and the more reliable the individual items, the more reliable the overall scale.

A questionnaire was completed by the raters after scoring of the recordings to assess if a debriefing could have been scored with the DMLES without observation of the simulation experience. There was agreement among the raters that each item was able to be scored without observation of the simulation since the purpose of scoring with the DMLES was to assess the debriefer, irrespective of scenario or level of participants. That is, rating each item as present or not present did not require knowledge of the objectives of the simulation.

Content validity is the degree to which a group of items represent an operational definition of a construct. Content validity was quantified through the content validity index (CVI), numerically representing the content relevance of instrument items using a 4-point rating scale, as rated by construct experts (Lynn, 1986). An acceptable level was determined as an index of .80 or greater. The item level CVI (I-CVI) of the DMLES was scored by three experts and subsequently calculated at 1.0 for 20 of the 33 items, indicating these 20 items were rated as highly relevant by all experts. The I-CVI of 8 of the 33 items was computed as greater than .80, while the I-CVI of 5 of the 33 items was computed as less than .80. The scale-level CVI (S-CVI) was computed by dividing the sum of each I-CVIs by the total number of items, yielding .92.

After scoring was completed, performance of each item was evaluated by entering the data from the 45 completed scales (15 recordings scored by each of the three raters) into Statistical Package for Social Scientists (SPSS). To assess the internal consistency of the scale, Cronbach's alpha was calculated. Cronbach's alpha has been widely used throughout the social sciences as an estimate of the internal consistency of a psychometric test. Internal consistency is considered excellent for a Cronbach's alpha greater than or equal to .9, between .8 and .9 is good consistency, and between .7 and .8 is an acceptable level of consistency. Cronbach's alpha for the total DMLES scale was .88 or acceptable. Cronbach's alpha for each element of the scale was as follows: engage (.39), explore (.51), explain (.73), elaborate (.79), evaluate (.77), and extend (.69). The low values of the subscales engage and explore are not unexpected, due to the low number of items in each subscale and the low number of observations.

One-way random effects intraclass correlation coefficients (ICC) were calculated for the total scale and each of the 33 items, to determine item reliability and interrater reliability (Shrout & Fleiss, 1979). The ICC is a descriptive statistic used when quantifying units organized into groups; the ratio of variance between groups to variance within those groups is used to determine interrater reliability. Values within the range of .40 to .75 were considered fair to good, while values greater than .75 are recommended in health research. The

ICC for the total scale was .86 ($p < 0.01$). The ICC for each phase of the DMLES was as follows: engage (.27), explore (.36), explain, (.69), elaborate (.78), evaluate (.73), and extend (.69). The low values of the subscales engage and explore were attributed to the removal of the items with zero variance, and can be interpreted theoretically as demonstrating a higher, thus acceptable, level of reliability.

In determining item inclusion or exclusion, consideration was first given to items that contributed the least amount of overall internal consistency (DeVellis, 2012). If all items were included in the final scale, the scale may be too long. However, consideration must be given to brevity versus reliability. For the final scale, five items were eliminated due to low CVI values, and two items were added to clarify behaviors specific to reflection-in-action and reflection-on-action were expanded for a final DMLES of 31 items.

The testing of the DMLES revealed several limitations. Testing of the scale across multiple sites would demonstrate if the psychometric properties of the scale could be replicated, as common to instrument validation. The small sample size and small number of recorded debriefing sessions may not be representative of all debriefers using the DML method. Further validity testing was indicated to ensure each item of the scale demonstrates the constructs represented by the DML method.

Debriefing for Meaningful Learning Inventory (DMLI)

The DML Inventory (DMLI) was developed as a subjective measure of a debriefer's DML debriefing, based on the 31 behavioral items of the DMLES.

Because each of the items of the DMLES describe an observable behavior, items that could be difficult to self-evaluate in a single statement were described in more than one item during DMLI development. As a result, the 31 items of the DMLES were expanded into 57 DMLI items for a self-reported measure.

Items one through 52 of the DMLI represented behaviors which should be consistently applied during DML debriefing, according to the original design and intent of the method. Each of the 52 items were scored with ordinal frequency options: *always, sometimes, never*. While the original DMLES items were scored as behaviors consistent with DML, which an observer could score as either *present* or *not present*, the subjective nature of the DMLI required these ordinal frequency options to allow respondents to reflect holistically on their typical debriefing behaviors. Additionally, ordinal frequency options allowed for responses that may have been difficult to limit to binary options.

Items 53 through 57 of the DMLI characterized the central concepts of DML: reflection-in-action, thinking like a nurse, reflection-on-action, challenging taken-for-granted assumptions, and reflection-beyond-action (Dreifuerst, 2012). Each of these five items was intended to assess whether or not a debriefer understood the central concepts of DML, and were therefore presented with binary options: *yes* or *no*.

Face validity of the DMLI was determined by the developer of DML and two experts in the method. Recommendations were made regarding wording of the items to increase clarity. Content validity was determined using the previously tested DMLES, since the content had not been changed from the original 31 DMLES items.

It was understood that data derived from self-report could have challenges related to validity and reliability, as responses could have been influenced by emotion, bias, and interpretation (Paulhus & Vazire, 2007). However, this methodology remains prevalent in the simulation literature (Berndt et al., 2015; Hawkins et al., 2014; Richardson, Goldsamt, Simmons, Gilmartin, & Jeffries, 2014). It is also recognized that the limitations of self-report are particularly evident when subjective measurement is associated with performance or competence, and when compared to objective measurement (Davis et al., 2006; Lai & Teng, 2011). However, in their meta-analysis on self-evaluations and performance outcomes, Zell and Krizan (2014) found an overall correlation (r = .29, SD = .11) between self-reported evaluation and objective performance. Given the challenge in recruiting participants for an observational study, it was determined that a self-report survey which evaluated both understanding and application of the central concepts of DML would provide useful data about the impact of training on faculty use of this particular debriefing method, and would be appropriate for this study.

Data collection. Survey respondents accessed the survey using Survey Monkey©, at which point they were asked if they debrief prelicensure baccalaureate nursing students. Those that indicated *no* were directed to an exit page, and were excluded from participating (n = 3). Those that indicated *yes* (*n* = 283) were prompted to choose whether or not to participate in the study by clicking *yes* or *no*.

Each participant then answered demographic questions, identified which debriefing method they used, and described the type of debriefing training they had received (Appendix E). Participants then answered the 52 items of the survey describing how they apply the central concepts of DML in simulation debriefing with prelicensure baccalaureate nursing students. The final five items of the survey were designed to assess understanding of the central concepts of DML.

1. Is the DMLI a valid measure of DML understanding and application?
2. Is there a difference in how many of the central concepts associated with DML that debriefers understood, when they were grouped according to the training they received?
3. Is there a difference in how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received?
4. Is there an interaction between how many of the central concepts associated with DML that debriefers understood, and how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received?

Hypotheses for the study were as follows:

$H1_0$: The DMLI is not a valid measure of DML understanding and application.

$H1_1$: The DMLI is a valid measure of DML understanding and application.

$H2_0$: There is no difference in how many of the central concepts associated with DML that debriefers understood, when they were grouped according to the training they received.

$H2_1$: There is a difference in how many of the central concepts associated with DML that debriefers understood, when they were grouped according to the training they received.

$H3_0$: There is no difference in how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received.

$H3_1$: There is a difference in how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received.

$H4_0$: There is no interaction between how many of the central concepts associated with DML that debriefers understood, and how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received.

$H4_1$: There is an interaction between how many of the central concepts associated with DML that debriefers understood, and how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received.

Data Analysis

After participants completed the DMLI using Survey Monkey©, survey response data were exported for data cleaning. Of the 276 participants who agreed to participate in the study, 242 completed the survey. Data from eight participants were removed from the data set due to incomplete surveys, yielding a sample of 234. Item responses were not forced; therefore, missing responses were accounted for by listwise deletion (McKnight, McKnight, Sidani, & Figueredo, 2007). The ordinal frequency options of *always*, *sometimes*, *never*, *yes*, and *no* were each converted to integers for data analysis. Reverse coding was implemented for 20 items.

Latent GOLD© 5.1 (2015) was used to perform a confirmatory latent class factor analysis (LCFA) to confirm item groupings in the DMLI based on the theoretical model of DML. Because traditional confirmatory factor analysis (CFA) is limited by requiring all variables to be continuous and normally distributed, LCFA was used to account for the discrete latent variables of the DMLI. LCFA is the most common model-based clustering method for discrete data, which combines elements of latent class (LC) analysis and standard factor analysis (FA) (Dean & Raftery, 2010; Magidson & Verumunt, 2004; Vermunt & Magidson, 2000). LCFA models assume 1) the latent variables are discrete, 2) the variables are dichotomous, ordinal, or nominal, and 3) conditional distributions are assumed to be binomial or multinomial. The DFactor model within Latent GOLD© was used to estimate the LC cluster models from the DMLI data (Vermunt & Magdison, 2005).

The current version of SPSS was used for data management and generation of statistics. Descriptive statistics were used to summarize the demographic and DMLI data. Inferential statistics using one-way analysis of variance (ANOVA) were used to compare differences in mean DMLI sums when grouped by types of training. Two-way ANOVA was used to determine and compare interactions between DMLI sums, and the simple main effects of group type and understanding of the central concepts of DML. Initial analyses were conducted to ensure there was no violation of the assumptions of normality and homogeneity. The effect sizes were calculated to evaluate the findings. Results were considered statistically significant if the *p*-value was less than 0.05.

Summary

This chapter describes the methodology, participants, and research questions used in this research study. Debriefers who received debriefing training and facilitate simulation debriefing with baccalaureate nursing students were solicited to complete the DMLI. The purpose of the DMLI was to describe the impact of training on how consistently the central concepts of the method were understood and applied, when compared to debriefers who had not been trained in DML. Data collection and analysis were also described in this chapter. Findings will be shared and explained in Chapters IV and V.

Chapter IV Findings

This study investigates the impact of training in Debriefing for Meaningful Learning© (DML) on how consistently debriefers trained in this method understand and apply its central concepts during simulation debriefing with baccalaureate nursing students, particularly when grouped according to the type of training they received. To achieve this, a new instrument, the Debriefing for Meaningful Learning Inventory (DMLI) was analyzed for the model of fit to the DML method. The types and sources of debriefing training, the understanding of the central concepts of DML, and the application of DML by current debriefers were analyzed using the DMLI.

This chapter presents the results of the data analysis for the four stated research questions: (1) Is the DMLI a valid measure of DML understanding and application? (2) Is there a difference in how many of the central concepts associated with DML that debriefers understood, when they were grouped according to the training they received? (3) Is there a difference in how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received? (4) Is there an interaction between how many of the central concepts associated with DML that debriefers understood, and how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received?

Descriptive Statistics

Demographic information was important to this study because it describes the types and sources of training reported; this demographic information described the variables of interest. The types and sources of training were reported by participants, in addition to other characteristics that described debriefing with prelicensure baccalaureate nursing students.

Types of Debriefing Training

One variable of interest of this study was debriefing training. Data describing this training were first obtained through participant responses to the survey demographic questions. To describe debriefing training, participants were presented with options that were similar to those reported in a previous debriefing survey (Waznonis, 2015). Participants were also presented with an option of *other,* and were then directed to describe their debriefing training in a text box if the options presented did not describe their training. The total group of participants (*N* = 234) was then categorized according to the type of debriefing method for which they reported having received training. Table 2 summarizes the types of debriefing training. Because DML was the debriefing method investigated in this study, participants who selected DML were assigned to the DML group (*n* = 71), while respondents who selected any method that was not DML were assigned to the non-DML group (*n* = 173).

Table 2

Type of Debriefing Training

Debriefing Method	*N*	%
Debriefing for Meaningful Learning	71	(30%)
Debriefing with Good Judgment	43	(16%)
NLN[a] 3 phase process	33	(12%)
Gather-Analyze-Summarize	18	(7%)
3D Model[b]	11	(4%)
Outcome-Present-State Model	0	(0%)
Other	73	(27%)
Combination of methods	24	
PEARLS[c]	17	
Plus-delta and advocacy-inquiry	5	
Self-developed method	5	
Chamberlain prepared model	3	
DEEP[d]	3	
Tanner Clinical Judgment Model	2	
Alpha, Delta, Gamma	2	
School/Program developed model	2	
Unknown	2	
Beard Model	1	
INACSL[e] Standards of debriefing	1	
MedSim Design	1	
MSR[f]	1	
None	1	
Pivec Model	1	
Reflection questions	1	
Talking with students	1	

[a]National League for Nursing
[b]3D = Defusing, Discovering, & Deepening
[c]PEARLS = Promoting Excellence and Reflective Learning in Simulation
[d]DEEP = Debriefing Engaged Experienced Practitioners
[e]INACSL = International Nursing Association for Clinical Simulation and Learning
[f]MSR = Israel Center for Medical Simulation

Sources of Debriefing Training

Recognizing that debriefing training is offered through a variety of sources, participants were asked to describe all applicable sources of debriefing training received. Participants were presented with options that were reported as sources of training in a previous debriefing survey (Waznonis, 2015). Table 3 summarizes the sources of training reported by the participants.

Table 3

Source of Debriefing Training

	N	%
Conference/Workshop	186	(80%)
Mentor	128	(55%)
Local/regional training	95	(41%)
NLN-SIRC[a] online courses	73	(31%)
INACSL[b] Webinar	80	(34%)
WISER[c] Center	10	(4%)
SSH-CHSE[d] program	39	(17%)
Vendor representative	34	(15%)
Vendor meeting	33	(14%)
Center for Medical Simulation (*Harvard*)	31	(13)
Center for Health Science Interprofessional Education, Research and Practice (University of Washington)	10	(4%)
Graduate level certificate program	14	(6%)
Michael S. Gordon Center for Research in Med. Education	1	(0.4%)
Other	62	(27%)
None	4	(2%)

[a]NLN-SIRC = National League for Nursing Simulation Innovation Resource Center
[b]INACSL = International Nursing Association for Clinical Simulation and Learning
[c]WISER = Peter M. Winter Institute for Simulation Education and Research
[d]SSH-CHSE = Society for Simulation in Healthcare Certified Healthcare Simulation Educator

To gain a deeper understanding of the sources specific to DML training, participants who identified receiving training in DML were asked additional questions to describe the specific DML training they received. Table 4 summarizes the characteristics of the training reported by the DML group.

Table 4

Training Characteristics of the DML Group

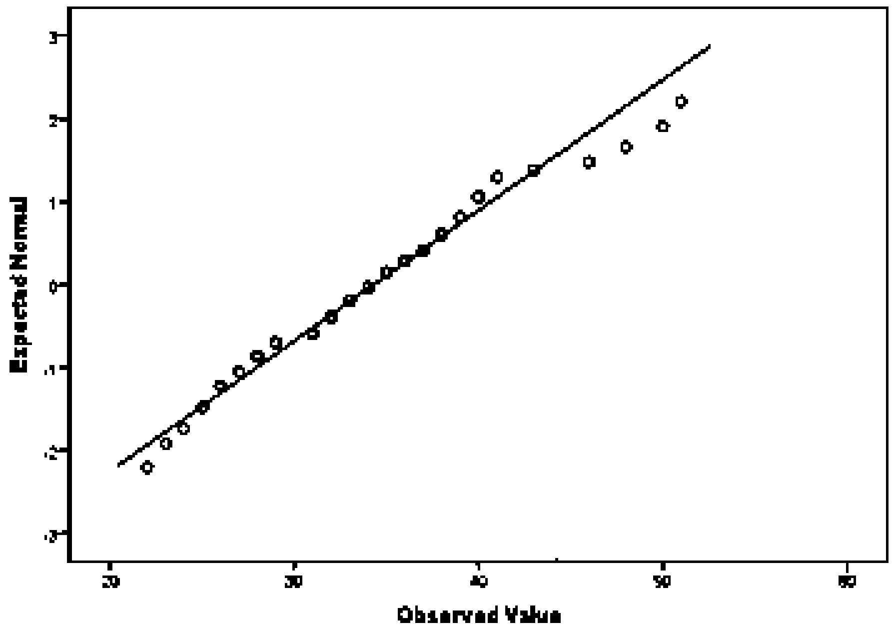
	N	%
Number of times used DML per semester[a]		
2-4 times	19	(27%)
5-8 times	13	(18%)
More than 8 times	39	(55%)
DML training[a]		
Read more than one article	46	(65%)
Watched a colleague use DML	36	(51%)
DML workshop/conference	35	(49%)
Read one article	5	(7%)
Attended a train-the-trainer session	18	(25%)
Year of training[b]		
2010	3	(1%)
2011	3	(1%)
2012	9	(23%)
2013	5	(13%)
2014	4	(10%)
2015	4	(10%)
2016	1	(0.3%)
Training included evaluation/feedback[d]		
Yes	24	(35%)
No	45	(65%)
Number of times DML has been evaluated since training[a]		
Never	30	(42%)
1-3 times	23	(32%)
4-6 times	5	(7%)
Greater than 6 times	7	(10%)
I believe I implement DML well		
Strongly agree	14	(21%)
Agree	45	(65%)
Disagree	9	(14%)
Strongly disagree	0	(0%)

[a]*n* = 71
[b]*n* = 40
[c]*n* = 41
[d]*n* = 69

Testing the Research Questions

Descriptive and inferential statistics were used to examine the four research questions guiding this study. Prior to data analysis for each of the research questions, the data were examined to ensure all assumptions for statistical tests were met. Analysis for this study is based on four assumptions that must be met to be accurate: (a) normally distributed data, (b) homogeneity of variance, (c) interval data, and (d) independence (Tabachnik & Fidell, 2013). DMLI scores were normally distributed for the DML group with a skewness of 0.395 (*SE* = .285) and kurtosis of 0.170 (*SE* = .563), and for the non-DML group with a skewness of -0.457 (*SE* = .190) and kurtosis of -0.302 (*SE* = .378). Visual inspection of Normal Q-Q Plots for the DML and the non-DML group also indicated the DMLI sums were normally distributed (Figure 1).

Q-Q Plot for DML group



Q-Q Plot for Non-DML Group

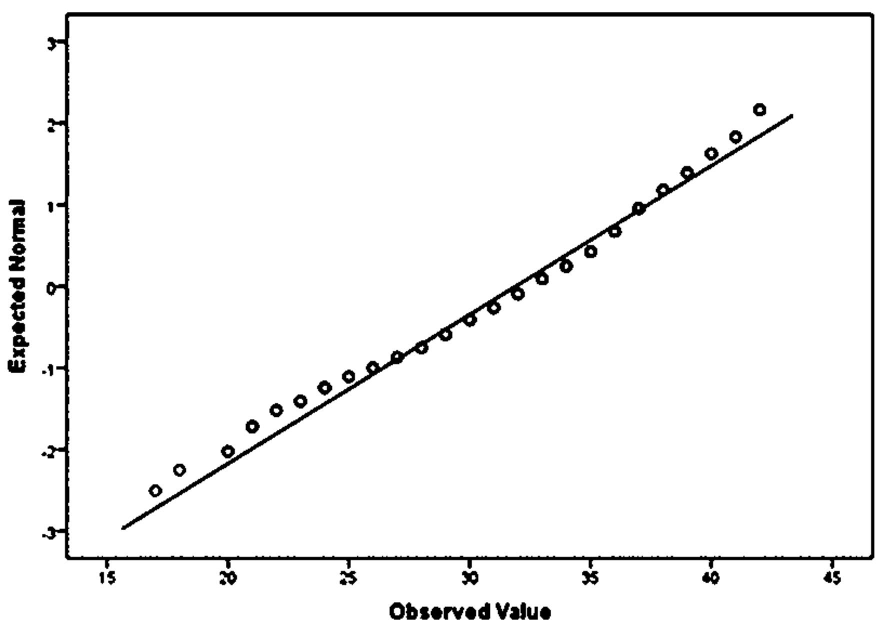


Figure 1. Q-Q Plots by Group Type

Research Question One

Question 1: Is the DMLI a valid measure of DML understanding and application? Face validity and content validity measures were used in addition to confirmatory DFactor analysis to answer this first question. Items one through 52 of the DMLI represented behaviors consistent with DML; participants self-reported how consistently they applied each of these behaviors with the ordinal frequency options of, *always, sometimes*, or *never*. Items 53 through 57 of the DMLI characterized the central concepts of DML; reflection-in-action, thinking like a nurse, reflection-on-action, challenging taken-for-granted assumptions, and reflection-beyond-action (Dreifuerst, 2012). Each of these five items was intended to assess whether or not a debriefer understood the central concepts of DML; participants self-reported their understanding with the binary options of *yes* or *no*. These items were not included in the factor analysis because they represented broader concepts that are embedded within each of the six E's of DML. Face validity of the DMLI was determined by the DML author and two experts in the method. Content validity was determined through previous testing of the DMLES, since the content had not varied from the original DMLES items.

Confirmatory factor analysis was used to confirm the DML application item groupings in the DMLI based on the theoretical model of DML. The goal of confirmatory factor analysis (CFA) is to confirm or reject the measurement theory (Bollen, 1989). However, CFA is limited by requiring all variables to be continuous and normally distributed. To account for the discrete variables generated through the DMLI data, latent class factor analysis (LCFA) was used.

LCFA is the most common model-based clustering method for discrete data, which combines elements of latent class (LC) analysis and standard factor analysis (FA) but is most closely analogous to cluster analysis (Dean & Raftery, 2010; Magidson & Vermunt, 2004; Vermunt & Magidson, 2000). LCFA models assume 1) the latent variables are discrete, 2) the variables are dichotomous, ordinal, or nominal, and 3) conditional distributions are assumed to be binomial or multinomial (Vermunt & Magidson, 2016). LCFA is similar to traditional FA in that they are both useful in data reduction. In addition, the latent classes, similar to factors, are unobserved constructs and are inferred from observed data (Lazarsfeld, Henry, & Anderson, 1968).

LCFA is a method used for identifying subtypes of related latent cases from data. LCFA uses data to estimate parameter values for the model. Each of the cases (the DMLI application items) were classified to the most likely latent class, similar to factors, by recruitment probabilities, which is the probability that for a randomly selected item of a given latent class, a given response pattern would be observed (Lazarsfeld et al.,1968; Vermunt & Magidson, 2005). Model fit is assessed by comparing cross-classification frequencies to expected frequencies predicted by the model; the difference is assessed with the L-squared statistic, which is similar to the chi-squared statistic. For this analysis of the DMLI application items, the bootstrapping approach was used, which requires no assumptions about the data as would chi-squared tests (Lazarsfeld et al.,1968).

The DFactor model, a type of LCFA, is a discrete latent trait model that was used to estimate the LC cluster models from the DMLI data with the Latent GOLD© software. The DFactor model accounts for latent variables called discrete factors (DFactors). A DFactor model is different from a LC cluster in that the DFactor model may contain more than one latent variable, and the categories are either ordered or dichotomous latent variables (Vermunt & Magidson, 2016).

DFactor analysis differs from traditional CFA in that the model is not linear, and solutions do not need to be rotated to be interpreted. One disadvantage of the DFactor model is that the parameters may be more difficult to interpret, such as the factor loading, factor-item correlations, factor correlations, and communalities. Because of this, a linear approximation of the maximum likelihood estimates is used, in order to provide similar output measures as traditional FA (Vermunt & Magidson, 2004).

LCFA was used to examine the inter-correlations that exist between the responses to the DMLI items, thus reducing the items into smaller groups known as factors (Tabachnik & Fidell, 2013). LCFA was used to evaluate if the items grouped according to the six E's of the DML method yield good fit. The six E's of DML include: engage, explore, explain, elaborate, evaluate, and extend.

The latent class approach to the DMLI data supports a six-class DFactor model, which provides an acceptable fit to the data, L-squared = 7.0803 with 85 degrees of freedom; p = 0.298. The Bayesian information criterion (BIC) statistic also indicated the six-class model was the preferred model (Table 5).

Table 5

Goodness of Fit Indices for Analysis with Structural Equation Models

X^2	L^2	BIC	AIC	CAIC	Bootstrap *p*-value
7.26	7.08	6630.79	6910.72	6545.79	0.298

Note. BIC, AIC, CAIC all based on L^2.

After selecting the number of factors in the model, the parameter estimates per factor were reviewed, as shown in Table 6. The parameters were checked for their significance of the estimate, indicating that the item can significantly discriminate (Table 7). The factors were ordered according to R-squared, which indicates how well the model predicts the DFactor score. R-squared for each DFactor are presented in Table 8. Each of the six DFactors were classified as the six E's of DML: engage, explore, explain, elaborate, evaluate, and extend.

Each of the 52 items were identified with keywords that describe their relevance to the respective DFactor. The keywords represent a DML behavior a debriefer would guide a student in, including a type of thinking, reasoning, or reflecting. A keyword may also represent the Socratic questioning used by a debriefer, or other behaviors consistent with DML. Each of the 52 items was reported in Table 7, although five items yielded an R-squared of zero. The R-squared of each item represents the amount of variance that is explained in the item by the factor. For instance, R-squared for Item 17, *Uncovering thinking*, is 0.3546. This means that 35% of the variance in Item 17 is explained by DFactor 4, *Elaborate*.

Data derived from the demographic section of the DMLI for the first research question suggest that indeed, the DMLI is a valid measure of DML. DFactor analysis of the DMLI yielded a model of good fit when the items were grouped together according to the 6 E's of DML. Therefore, the null hypothesis was rejected, and the alternative hypothesis was accepted.

Table 6

Parameter Estimates per DFactor

DFactor	Parameter Estimate	Wald Statistic	*p*-value
1	1.4778	5.0692	0.024
2	4.2848	5.2114	0.022
3	2.3974	6.275	0.012
4	3.3553	7.6418	0.006
5	1.9086	4.4002	0.036
6	0.3105	6.9408	0.009

Table 7

Pattern Matrix for DFactor Loadings

		DFactor						
Item	Description	DFactor 1 Engage	DFactor 2 Explore	DFactor 3 Explain	DFactor 4 Elaborate	DFactor 5 Evaluate	DFactor 6 Extend	R^2
1	Silently reflecting	-0.2529	0	0	0	0	0	0.064
2	Reconstructing	0	0	0	-0.1287	0	0	0.0166
3	Non-judging	0	0	0	0	0	0	0
4	Summarizing	-0.1636	0	0	0	0	0	0.0268
5	Socratic questioning	-0.3257	0	0	0	0	0	0.1061
6	Decision-making	0	0	-0.1697	0	0	0	0.0288
7	Summarizing	0	0	0	0	0	0	0
8	Generalizing	0	0	-0.1697	0	0	0	0.0766
9	Reflecting beyond	0	0	0	0	0	0.4879	0.238
10	Listening	0	0	0	0	0	0	0
11	Interpreting	0	-0.0317	0	0	0	0	0.001
12	Engaging	-0.1873	0	0	0	0	0	0.0351
13	Reflecting	0	0	0	0	-0.0157	0	0.0002
14	Reframing	0	0	0	0	0.235	0	0.0552
15	Clinical decisions	0	0	-0.363	0	0	0	0.1318
16	Reflecting	0	0	0	0	-0.1148	0	0.0132
17	Uncovering thinking	0	0	0	-0.5955	0	0	0.3546
18	Self-evaluating	0	0	0	0.1707	0	0	0.0292
19	Reflecting	0	0	0	0	-0.0982	0	0.0097
20	Clinical decisions	0	0	0	-0.4942	0	0	0.2442
21	Thinking & actions	0	0.8048	0	0	0	0	0.141

22	Socratic questioning	0.4225	0	0	0	0	0	0.1785
23	Participating	-0.2042	0	0	0	0	0	0.0417
24	Socratic questioning	0	-0.5099	0	0	0	0	0.26
25	Worksheets	-0.0673	0	0	0	0	0	0.0045
26	Decision-making	0	0	-0.5852	0	0	0	0.3424
27	Reflecting	0	0	0	-0.7128	0	0	0.5081
28	Reflecting	0	0	0	0	0.0645	0	0.0042
29	Assessments	0	0	-0.5441	0	0	0	0.296
30	Thinking & actions	0	0	-0.7446	0	0	0	0.5544
31	Thinking like a nurse	0	0	0	-0.0035	0	0	0
32	Collaborating	0	0	0	0	-0.0443	0	0.002
33	Reflecting beyond	0	0	0	0	0	0.7905	0.6249
34	Summarizing	0.216	0	0	0	0	0	0.0467
35	Actioned decisions	0	0	-0.4334	0	0	0	0.1879
36	Worksheets	0.1411	0	0	0	0	0	0.0199
37	Listening	-0.1336	0	0	0	0	0	0.0178
38	Reflecting	0	0	0	-0.3607	0	0	0.1301
39	Reflecting beyond	0	0	0	0	0	0.5434	0.2953
40	Clinical decisions	0	0	0	0	-0.1656	0	0.0274
41	Listening	-0.0064	0	0	0	0	0	0
42	Analyzing	0	0	0	0	0.23	0	0.0005
43	Anticipating	0	0	0	0	0	0.1931	0.0373
44	Decision-making	0	0	0	0	-0.5895	0	0.3475
45	Contributing factors	0	-0.7747	0	0	0	0	0.6002
46	Building explanations	0	0	0	-0.703	0	0	0.4942
47	Patient’s responses	0	0	-0.5957	0	0	0	0.3549
48	Parking emotions	-0.2025	0	0	0	0	0	0.041
49	Applying knowledge	0	0.3397	0	0	0	0	0.1154
50	Thinking & actions	0	-0.6086	0	0	0	0	0.3704

51	Thinking routines	0	-0.3082	0	0	0	0	0.095
52	Assimilating	0	0	0	0	0	0.1862	0.0347

Table 8

R-squared for DFactors

DFactor		R^2
DFactor 1	Engage	0.7841
DFactor 2	Explore	0.8762
DFactor 3	Explain	0.8984
DFactor 4	Elaborate	0.9117
DFactor 5	Evaluate	0.8308
DFactor 6	Extend	0.7690

Research Question Two

Question 2: *Is there a difference in how many of the central concepts associated with DML that debriefers understood, when they were grouped according to the training they received?* The second research question was examined with descriptive statistical analysis of items 52 through 57. A one-way ANOVA was used to compare the amount of variance between the independent variable of type of debriefing training, and the dependent variable of DMLI sum. While determining reliability coefficients was considered, this was not included in the analysis because the DML central concepts are different domains of a complex construct, and would not have provided additional useful information.

The total counts of central concepts understood were compared for items 53 through 57. The DML central concepts of understanding (CCU) are represented in items 53 through 57 of the DMLI. Responses to these items demonstrate a respondent's understanding of the central concepts of DML: reflection-in-action (CCU-A), thinking like a nurse (CCU-B), reflection-on-action (CCU-C), challenging taken for granted assumptions (CCU-D), and reflection-beyond-action (CCU-E) (Dreifuerst, 2012). Counts of correct responses for each CCU were calculated for each respondent and analyzed (Table 9).

Table 9

Understanding of CCU by Group

	DML group		Non-DML group		Total group	
	Counts	Frequency	Counts	Frequency	Counts	Frequency
CCU-A	35 (*n*=71)	49.3%	94 (*n*=159)	59.1%	129 (*n*=230)	56.1%
CCU-B	69 (*n*=70)	98.6%	140 (*n*=160)	87.5%	209 (*n*=230)	90.9%
CCU-C	20 (*n*=71)	28.2%	47 (*n*=159)	29.6%	67 (*n*=230)	29.1%
CCU-D	65 (*n*=70)	92.9%	128 (*n*=160)	80.0%	193 (*n*=230)	83.9%
CCU-E	19 (*n*=71)	26.8%	55 (*n*=161)	34.2%	74 (*n*=232)	31.9%

As illustrated in Table 10, there were statistically significant differences between groups for understanding of CCU-B (thinking like a nurse), $F(1,228) = 7.362$, $p = .007$, $n^2 = .031$, and CCU-D (challenging taken-for-granted assumptions), $F(1,230) = 1.238$, $p = .015$, $n^2 = .026$. There was an effect of group type on the understanding of the DML central concepts CCU-B (thinking like a nurse) and CCU-D (challenging taken-for-granted assumptions), although the effect size was small. The null hypothesis was rejected and the alternative hypothesis was accepted, as there were differences in DMLI scores of participants when grouped according to the type of training received.

Table 10

CCU Significance Levels

CCU	Measure	Sum of Squares	df	Mean Square	*F*	Sig.[a]	n^2
CCU-A (Reflection-in-Action)							
	Between Groups	.474	1	.474	1.923	.167	0.008
	Within Groups	56.174	228	.246			
	Total	56.648	229				
CCU-B (Thinking Like a Nurse)							
	Between Groups	.597	1	.597	7.362	.007	0.031
	Within Groups	18.486	228	.081			
	Total	19.083	229				
CCU-C (Reflection-on-Action)							
	Between Groups	.009	1	.009	.046	.831	0.000
	Within Groups	47.473	228	.208			
	Total	47.483	229				
CCU-D (Challenging Taken-for-granted Assumptions)							
	Between Groups	.805	1	.805	6.069	.015	0.026
	Within Groups	30.243	228	.133			
	Total	31.048	229				
CCU-E (Reflection-beyond-Action)							
	Between Groups	.270	1	.270	1.238	.267	0.005
	Within Groups	50.127	230	.218			
	Total	50.397	231				

[a]Computed using alpha = .05

Research Question Three

Question 3: Is there a difference in how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received? To answer the third question, one-way analysis of variance (ANOVA) was used to compare the amount of variance between the independent variable, type of debriefing training, and the dependent variable, DMLI sum. Mean DMLI sum differences between groups were computed with alpha set at .05. Eta squared was computed to determine effect size and post hoc tests were conducted if there was a statistically significant difference. Although ANOVA is considered to be robust to moderate departures from the homogeneity of variance assumption, Levene's Test of Equality of Variances was used to check the assumption that the variances of the two groups were not significantly different. The Levene's statistic indicates that group variances can be treated as equal if the significance value is greater than 0.05.

The dependent variable, DMLI sum, was calculated by adding the correct responses of items one through 52. The DMLI sum represents how consistently respondents describe applying the process of DML. The correct items of the DMLI were summed for each respondent. Mean scores of each group were calculated and analyzed to answer the research questions.

Respondents indicated their application of the DML behavior described in each item with the options: *always*, *sometimes*, or *never*. All responses were coded numerically as either consistent with DML (1), or not consistent with DML (0). The total sum of the correct responses of DMLI items one through 52 were calculated for each respondent. The mean DMLI sum was then calculated for the DML and the non-DML group. Table 11 illustrates mean DMLI sums for both groups. The DML group scored a higher number of items that are consistent with DML behaviors (M = 34.31, SD = 6.360) than the non-DML group (M = 31.90, SD = 5.475).

Table 11

Mean DMLI Sum by Group

	N	*M*	*Mdn*	*SD*	Min	Max	Range	Variance
DML group	71	34.31	34.00	6.360	22	51	29	40.445
Non-DML group	163	31.90	32.00	5.475	17	42	25	29.978
Total	234	32.63	33.00	5.850	17	51	34	34.225

A one-way ANOVA was performed to compare the amount of variance of mean DMLI scores between the DML group and the non-DML group. Levene's Test for Equality of Variances was conducted (Table 12) prior to analysis, showing that group variances can be treated as equal, $F(1,232) = 1.267$, $p = .262$ (Table 12).

Table 12

Test for Equality of Variances (Levene's)

F	df1	df2	*p*
1.267	1	232	0.262

The mean DMLI sums were statistically significantly different between groups, $F(1, 232) = 8.655$, $p = .004$, $n^2 = 0.036$ (Table 13). The mean increase (2.408) from the non-DML group to the DML group (2.408, 95% CI[.795,4.021]) was also statistically significant ($p = .004$), as shown in Table 14. Pairwise comparisons (Table 13) revealed that the mean increase from the non-DML group ($M = 31.90$, $SD = 5.475$) to the DML group ($M = 34.31$, $SD = 6.360$) was statistically significant ($p = .004$).

Table 13

One-way ANOVA of DMLI Sum

		Sum of Squares	df	Mean Squares	F	Sig	n^2
Sum*Group	Between Groups	286.781	1	286.781	8.655	.004	0.036
	Within Group	7687.613	232	33.136			
	Total	7974.393	233				

Table 14

Pairwise Comparisons

		Mean Difference	*SE*	Sig	95% CI LL	95% CI UL
GROUP	GROUP					
DML	Non-DML	2.408*	.819	.004	.795	4.021
Non-DML	DML	-2.408*	.819	.004	-4.021	-.795

Note. Dependent variable: DMLI Sum. CI = confidence interval; LL = lower limit, UL = upper limit
* The mean difference is significant at the .05 level.

The mean DMLI sums were compared between the two groups according to highest academic agree achieved by participants (Table 15). Within the DML group, a pattern was noted of increasing mean DMLI sums with each advanced degree; doctoral prepared participants in the DML group demonstrated the highest mean DMLI sum.

Table 15

DMLI Sum by Type of Training (Group) and Highest Academic Degree

Degree	DML			Non-DML		
	N	*M*	*SD*	*N*	*M*	*SD*
Bachelor	5	31.40	7.503	16	30.81	4.636
Master	47	32.60	5.570	92	32.03	5.918
Doctorate	19	39.32	5.386	55	32.00	4.955

A two-way ANOVA was conducted to examine the interaction between the group type, DML and non-DML groups, and the source of training on mean DMLI sums (Table 16). Residual analysis was performed to test for the assumptions of the two-way ANOVA. Homogeneity of variances was assessed by Levene's test. There were no outliers, residuals were normally distributed (p >.05) and there was homogeneity of variances for analyses except for the source of training of vendor meeting (p = .001).

There were statistically significant interactions between group type and the training sources of vendor representative (F[1,226] = 4.902, p = .028, n^2 = .021), vendor meeting (F[1,226] = 4.417, p = .037, n^2 = .019), NLN-SIRC online courses (F[1,226] = 7.398, p = .007, n^2 = .031), and CMS (F[1,226] = 9.483, p = .002, n^2 = .040).

The mean DMLI sums within the DML group were further examined to understand differences related to sources of DML training. Patterson (2009) reported literature, conferences, and colleagues as sources of evidence nurse educators use to build teaching practice. Table 17 lists the sources of training selected by participants in the DML group. Participants could select all applicable sources of training received. The highest mean DMLI sum was within the group of participants who received training for the National Council of State Boards of Nursing (NCSBN) for participation in the National Simulation Study (NSS) (M = 38.75, SD = 2.630, n = 4). Second to this group were participants who received training at a DML workshop or conference (M = 36.57, SD = 6.621, n = 35). This mean, however, was only higher by .01 than those who reported attending a train-the-trainer DML training (M = 36.56, SD = 6.224, n = 18). The lowest mean DMLI sum was among those who reported training solely through reading one DML article (M = 31.00, SD = 5.099, n = 5).

Table 16

DMLI Sum by Type of Training (Group) and Source of Training

	DML			Non-DML			Interaction Group*Source				
	N	*M*	*SD*	*N*	*M*	*SD*	Levene's	*F*	Sig.[a]	n^2	Observed Power
CMS (Harvard)	3	45.00	7.000	28	32.14	5.563	.904	9.483	.002	.040	.866
Vendor Meeting	13	38.69	8.300	20	32.55	5.073	.183	4.417	.037	.019	.553
Vendor Rep	13	38.54	9.404	21	32.19	5.335	.001	4.902	.028	.021	.597
NLN–SIRC courses	20	38.35	7.329	53	32.49	4.925	.104	7.398	.007	.031	.772
WISER Center	4	37.00	3.742	6	34.17	5.345	.417	.017	.896	.000	.052
Grad certificate program	2	37.00	1.414	12	33.25	5.910	.278	.086	.770	.000	.060
SSH-CHSE Program	11	36.00	7.642	28	31.75	5.726	.415	.951	.330	.004	.163
INACSL Webinar	19	35.63	6.914	61	32.10	5.353	.703	.682	.410	.003	.130
Center for HSIERP	2	34.50	9.192	8	33.38	4.406	.456	.085	.771	.000	.060
Local/Regional Meeting	30	34.40	6.021	65	33.31	4.978	.467	1.759	.186	.008	.262
Conference/Workshop	58	34.26	6.523	128	32.21	5.252	.377	.683	.410	.003	.130
Mentor	32	33.72	6.055	96	31.80	5.453	.675	.253	.615	.001	.079
Michael S Gordon Center	0	-	-	1	41.00	-	.210	-	-	.000	-
Other	14	32.29	5.744	48	32.75	4.844	.154	3.556	.061	.015	.467
None	1	29.00	-	3	29.00	3.000	.209	.131	.717	.001	.236

[a]Computed using alpha = .05

Table 17

DMLI Sum by Source of DML Training

Source of Training	*N*	*M*	*SD*	*SE*
NCSBN Training for NSS	4	38.75	2.630	1.014
DML Workshop/Conference	58	36.57	6.621	1.119
Attended train-the-trainer session	18	36.56	6.224	1.467
Read more than one article	46	35.63	6.482	.956
Watched a colleague use DML	36	34.31	6.360	.755
Read one article	5	31.00	5.099	2.280

Multiple sources of training were reported by participants within the DML group (*n* = 56). Four commonly reported combinations of sources of training included DML workshop or conference attendance. Each combination demonstrated increasing mean DMLI sums with each additional source of training (Table 18). Reading one article in addition to attending a DML workshop or conference resulted in a higher mean score (*M* = 38.00, *SD* = 6.97, *n* = 24) than just attending a workshop or conference alone (*M* = 36.57, *SD* = 6.621, *n* = 35). Watching a colleague use DML in addition to DML workshop or conference attendance resulted in yet a higher mean DMLI sum (*M* = 38.53, *SD* = 8.62, *n* = 15). Furthermore, both watching a colleague and reading more than one article added to DML workshop or conference attendance yielded an even higher mean DMLI sum (*M* = 40.67, *SD* = 8.14, *n* = 12). The addition of a fourth source of training to this combination, attending a train-the-trainer session, resulted in the highest mean score of the combined sources of training (*M* = 41.80, *SD* = 7.19, *n* = 5).

Table 18

DMLI Sum by Combined Sources of Training

Training	*N*	*M*	*SD*	*SE*
DML workshop/Conference and Read more than one article	24	38.00	6.97	1.93
DML Workshop/Conference and Watched a colleague use DML	15	38.53	8.62	3.07
DML Workshop/Conference and Watched a colleague use DML and Read more than one article	12	40.67	8.14	2.35
DML Workshop/Conference and Read more than one article and Watched a colleague use DML and Attended a train-the-trainer session	5	41.80	7.19	3.22

Participants were asked to describe how well they believe they implement DML with baccalaureate nursing students by selecting one of four options: *strongly agree*, *agree*, *disagree*, or *strongly disagree*. No participants selected *strongly disagree*. Mean DMLI sums increased as self-perception of implementation also increased, as illustrated in Table 19.

Table 19

DMLI Sum by Self-Perception "I believe I implement DML well"

Response	*N*	*M*	*SD*	*SE*
Strongly Agree	14	37.57	4.636	.597
Agree	45	33.93	6.770	.354
Disagree	9	31.56	5.270	.717
Strongly Disagree	0	0	0	0

Participants were also asked if their initial DML training included an evaluation of their DML debriefing; no evaluation was reported by 56.3% (n = 23), while evaluation was reported by 32.4% (n = 23). Frequency of debriefing evaluation since the initial DML training is reported in Table 20.

Table 20

Debriefing Evaluation after DML Training

Evaluation	Frequency	Percent
Never	30	42.3
0-3 times	23	32.4
4-6 times	5	7.0
>6 times	7	9.9

In conclusion, the null hypothesis for the third research question was rejected, and the alternative hypothesis was accepted indicating that there were differences in mean DMLI sums of when participants were grouped according to the type of training received. Furthermore, there were differences in mean DMLI sums in the DML group according to the source of DML training received, and according to additional training received.

Research Question Four

Question 4: Is there an interaction between how many of the central concepts associated with DML that debriefers understood, and how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received? The fourth research question was examined by conducting a two-way between-subjects ANOVA to determine the interaction of the independent and dependent variables, and the significance of the main effects. The independent variables were the group type and the five individual measurements of understanding of the central concepts (CCU) of DML. The dependent variable was the mean DMLI sum. A Bonferroni adjustment was applied, interaction and main interactions effects were computed, and mean differences were examined. Alpha was set at .05 and effect size was computed.

If there was a statistically significant ordinal interaction, the main effects were analyzed and reported (Maxwell & Delaney, 2004). If an interaction effect was not statistically significant, the main effects were analyzed and reported (Faraway, 2015). If a main effect was statistically significant, a post hoc analysis (e.g., all pairwise comparisons) (Howell, 2010) was also conducted. If there was no interaction effect, the main effects are equal and can be considered together as one, requiring the reporting of the effect on the dependent variable. A non-statistically significant interaction does not indicate that an interaction effect does not exist within the population (Faraway, 2015; Fox, 2008; Searle, 2006).

Residual analysis was performed to test for the assumptions of the two-way ANOVA. DMLI sums were normally distributed for the DML group with a skewness of .395 (*SE* = .285) and kurtosis of .170 (*SE* = .563) and for the non-DML group with a skewness of -.457 (*SE* = .190) and kurtosis of -.302 (*SE* = .378). The Shapiro-Wilk's test showed that DMLI sums were normally distributed for the DML group ($p < .05$), but not for the non-DML group. Because of the larger sample size (> 100) of the non-DML group ($n = 163$), and because the skewness and kurtosis were not statistically significant (±2.58), the data were not transformed and the ANOVA was considered robust to this deviation from normality (Maxwell & Delaney, 2004; Tabachnick & Fidell, 2013). Prior to the two-way ANOVA for each CCU, homogeneity of variances was determined by Levene's Test for Equality of Variances (Table 21).

Table 21

Test for Equality of Variances (Levene's) for CCUs

CCU	*F*	df1	df2	*Sig.*[a]
CCU-A (Reflection-in-Action)	2.565	3	226	0.055
CCU-B (Thinking Like a Nurse)	0.946	3	226	0.419
CCU-C (Reflection-on-Action)	1.469	3	226	0.224
CCU-D (Challenging Taken-for-granted Assumptions)	0.911	3	226	0.436
CCU-E (Reflection-beyond-Action)	0.342	3	226	0.795

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
Design: Intercept + GROUP + CCU + GROUP*CCU
[a]Computed using alpha = .05

CCU-A (Reflection-in-Action). A two-way ANOVA was conducted to determine if there was an interaction between the independent variables of group type and understanding of CCU-A, reflection-in-action, on DMLI scores. The interaction effect between group type and CCU-A on DMLI sums was not statistically significant, $F(1,226) = 3.834$, $p = .051$, partial $n^2 = .018$. Although the interaction was not statistically significant, this does not mean that an interaction effect does not exist in the population (Faraway, 2016; Searle, 2006). A clustered bar chart (Figure 2) depicts the differences in the dependent variable (mean DMLI sums), based on the two independent variables (group type and CCU-A).

The main effects were analyzed next using Type III sums of squares to determine if an interaction effect might actually exist in the population (Fox, 2008; Maxwell & Delaney, 2004). All univariate tests and pairwise comparisons were run between the cells of the simple main effects, and a Bonferroni adjustment was made to correct for multiple comparisons within each simple main effect separately. The main effect of group type on mean DMLI score was statistically significant, $F(1,226) = 9.262$, $p = .003$, $n^2 = .039$, specifically within the DML group, $F(1,226) = 9.301$, $p = .003$, $n^2 = .040$ (Table 22). The main effect of the variable CCU-A on the mean DMLI score was statistically significantly different, $F(1,226) = 9.528$, $p = .002$, partial $n^2 = .040$, specifically those that did not indicate understanding CCU-A, $F(1,226) = 13.138$, $p = .000$, partial $n^2 = .055$. Although the effect size was small for both main effects, both DML training and the central concept represented in CCU-A (reflection-in-action) did have a statistically significant increase on the mean DMLI sum despite the absence of a statistically significant interaction.

Table 22

Test of Between Subjects Effects for Group Type and CCU-A (Reflection-in-Action)

Source	Type III Sum of Squares	df	Mean Square	*F*	Sig.[b]	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	594.227[a]	3	198.076	6.279	.000	.077	18.838	.964
Intercept	212966.720	1	212966.720	6751.502	.000	.968	6751.502	1.000
GROUP	292.171	1	292.171	9.262	.003	.039	9.262	.858
CCU-A	300.537	1	300.537	9.528	.002	.040	9.528	.867
GROUP*CCU-A	120.935	1	120.935	3.834	.051	.017	3.834	.496
Error	7128.855	226	31.544					
Total	253529.000	230						
Corrected Total	7723.083	229						

[a]R Squared = .077 (Adjusted R Squared = .065).
[b]Computed using alpha = .05

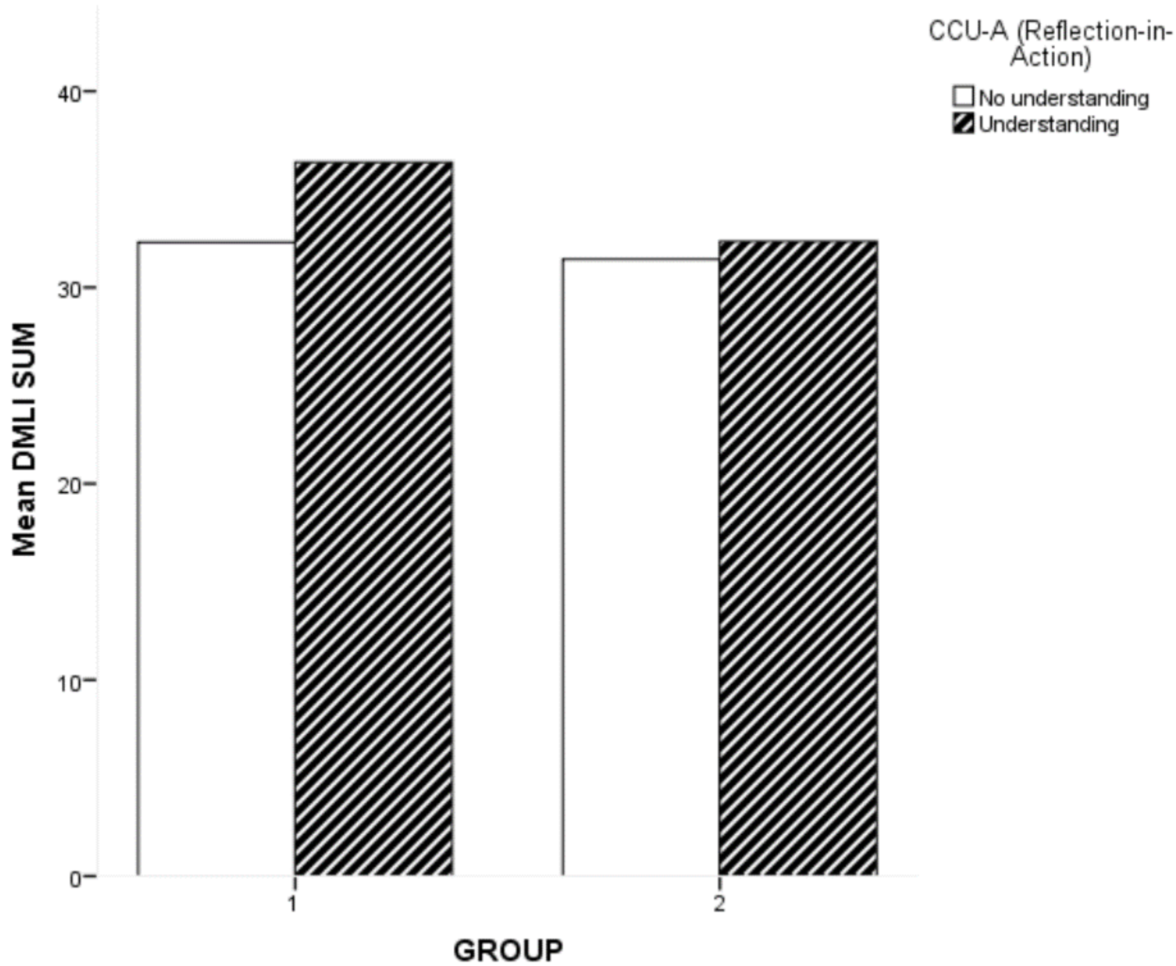


Figure 2. Two-way ANOVA (Group*CCU-A). Clustered Bar Chart with Error Bars: 95% CI. 1 = DML Group, 2 = Non-DML Group

As presented in Table 23, the mean DMLI sum for those who indicated understanding of CCU-A, reflection-in-action, was highest within the DML trained group (*M* = 36.371, *SE* = .949). There was a statistically significant difference, $F(1,226) = 9.301$, $p = .003$, partial $n^2 = .040$ in the mean DMLI difference between those who did and did not indicate understanding of CCU-A (Table 24). Of those who were trained in DML, the mean DMLI difference was higher by

4.066, 95% CI[1.439, 6.693] for those who indicated understanding CCU-A than those who did not (Table 25).

Table 23

Mean DMLI Sum by Group Type and CCU-A (Reflection-in-Action) Response

Group	CCU-A	*N*	*M*	SD	SE	95% CI	
						LL	UL
DML							
	No understanding	36	32.31	4.67	.94	30.46	34.15
	Understanding	35	36.37	7.22	.96	34.50	38.24
Non-DML							
	No understanding	65	31.43	5.40	.70	30.06	32.80
	Understanding	94	32.34	5.41	.58	31.99	33.48

Note. CI = confidence interval; LL = lower limit, UL = upper limit

Table 24

Univariate Tests for Group Type and CCU-A (Reflection-in-Action)

Simple Main Effects	Sum of Squares	df	Mean Square	*F*	Sig.[a]	Partial Eta Squared	Observed Power[a]
Group Type							
Contrast	292.171	1	292.171	9.262	.003	.039	.858
Error	7128.855	226	31.544				
DML Group							
Contrast	293.373	1	293.373	9.301	.003	.040	.859
Error	7128.855	226	31.544				
Non-DML Group							
Contrast	31.798	1	31.798	1.008	.316	.040	.859
Error	7128.855	226	31.544				
CCU-A							
Contrast	300.537	1	300.537	9.528	.002	.040	.867
Error	7128.855	226	31.544				
No Understanding of CCU-A							
Contrast	414.412	1	414.412	13.138	.000	.055	.950
Error	7128.855	226	31.544				
Understanding of CCU-A							
Contrast	17.730	1	17.730	.562	.454	.002	.116
Error	7128.855	226	31.544				

Note. Dependent variable: DMLI Sum.
[a]Computed using alpha = .05

Table 25

Pairwise Comparisons for Group Type and CCU-A (Reflection-in-Action)

Group	CCU-B		Mean Difference	*SE*	Sig.[a]	95% CI for Difference[b] LL	UL
DML	0	1	-4.066*	1.33	.003	-6.693	-1.439
	1	0	4.066*	1.33	.003	1.439	6.693
Non-DML							
	0	1	-.910	.906	.316	-2.695	.876
	1	0	.910	.906	.316	-.876	2.695

Note. Dependent variable: DMLI Sum. * The mean difference is significant at the .05 level.
CI = confidence interval; LL = lower limit, UL = upper limit
0 = No understanding of CCU-A,1 = Understanding of CCU-A.
[a]Computed using alpha = .05
[a]Adjustment for multiple comparisons: Bonferroni.

CCU-B (Thinking Like a Nurse). A two-way ANOVA was conducted to determine if there was an interaction between the two independent variables of group type and CCU-B, thinking like a nurse, on mean DMLI sums. As presented in Table 26, there was a statistically significant interaction between group type and CCU-B on mean DMLI sum, $F(1, 226) = 4.172$, $p = .042$, partial $n^2 = .018$. Because of the ordinal and disordinal interactions, analyses of the simple main effects for CCU-B and group type were performed with a Bonferroni adjustment. A clustered bar chart (Figure 3) depicts the differences in the dependent variable (mean DMLI sums), based on the two independent variables (group type and CCU-B).

The main effects were analyzed next using Type III sums of squares to determine if an interaction effect might exist in the population (Fox, 2008; Maxwell & Delaney, 2004). All univariate tests and pairwise comparisons were run between the cells of the simple main effects, and a Bonferroni adjustment was made to correct for multiple comparisons within each simple main effect separately. As shown in Table 27, CCU-B had a statistically significant difference in mean DMLI sum for those who indicated understanding of CCU-B, $F(1,226) = 9.418$, $p = .002$, partial $n^2 = .040$. There was a statistically significant difference in mean DMLI sums for the DML group, $F(1,226) = 4.905$, $p = .028$, partial $n^2 = .021$. Although the effect size was small to moderate for the simple main effects, both group type and understanding of the central concept represented in CCU-B did have a statistically significant increase on DMLI sums.

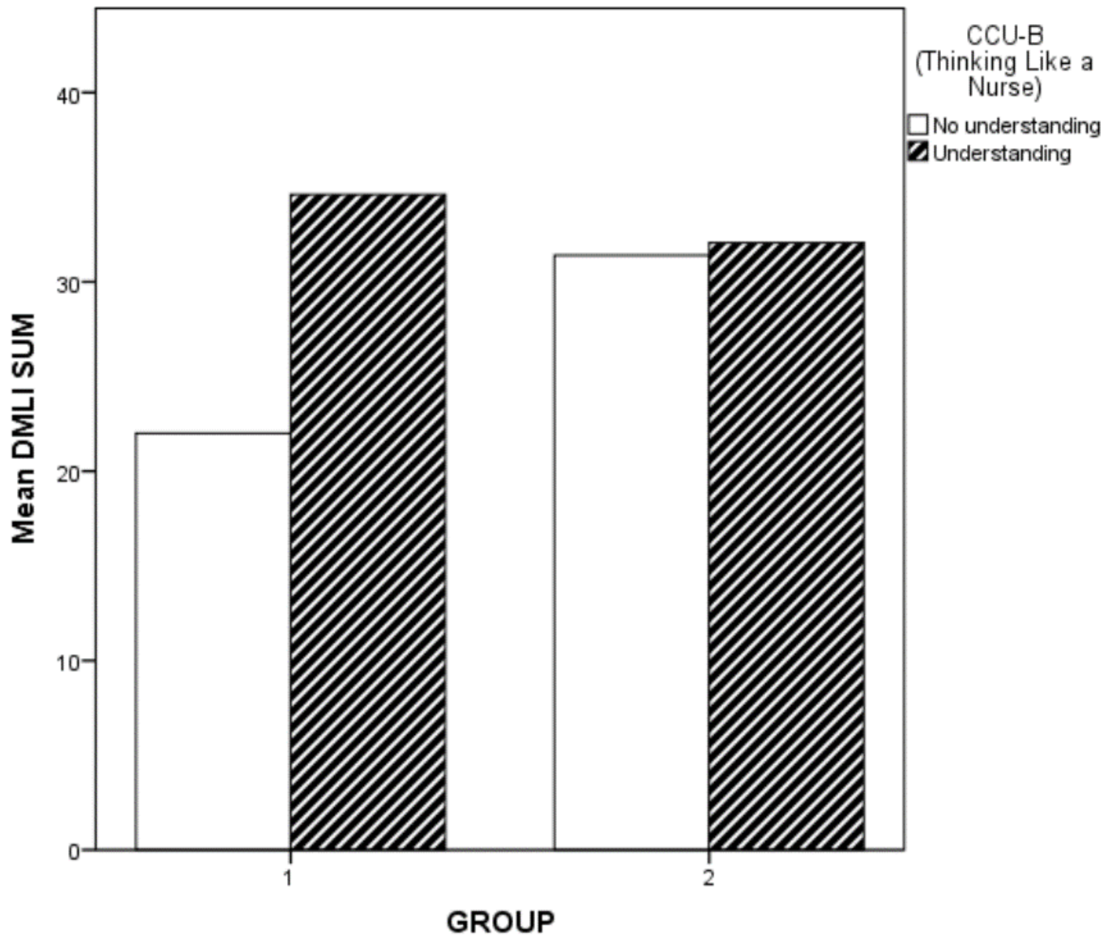


Figure 3. Two-way ANOVA (Group*CCU-B). Clustered Bar Chart with Error Bars: 95% CI. 1 = DML Group, 2 = Non-DML Group

Table 26

Test of Between Subjects Effects for Group Type and CCU-B (Thinking Like a Nurse)

Source	Type III Sum of Squares	df	Mean Square	*F*	Sig.[b]	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	457.414[a]	3	152.471	4.772	.003	.060	14.316	.898
Intercept	13452.152	1	13452.152	421.033	.000	.651	421.033	1.000
GROUP	43.766	1	43.766	1.370	.243	.006	1.370	.214
CCU-B	164.219	1	164.219	5.140	.024	.022	5.140	.617
GROUP*CCU-B	133.291	1	133.291	4.172	.042	.018	4.172	.529
Error	7220.778	226	31.950					
Total	253942.000	230						
Corrected Total	7678.191	229						

[a]R Squared = .060 (Adjusted R Squared = .047).
[b]Computed using alpha = .05

Table 27

Univariate Tests for Group Type and CCU-B (Thinking Like a Nurse)

Simple Main Effects	Sum of Squares	df	Mean Square	*F*	Sig.[a]	Partial Eta Squared	Observed Power[a]
Group type							
Contrast	43.766	1	43.766	1.370	.243	.006	.214
Error	7220.778	226	31.950				
DML Group							
Contrast	156.708	1	156.708	4.905	.028	.021	.597
Error	7220.778	226	31.950				
Non-DML Group							
Contrast	7.557	1	7.557	.237	.627	.001	.077
Error	7220.778	226	31.950				
CCU-B							
Contrast	164.219	1	164.219	5.140	.024	.022	.617
Error	7220.778	226	31.950				
No Understanding of CCU-B							
Contrast	84.152	1	84.152	2.634	.106	.012	.366
Error	7220.778	226	31.950				
Understanding of CCU-B							
Contrast	300.912	1	300.912	9.418	.002	.040	.863
Error	7128.855	226	31.544				

Note. Dependent variable: DMLI Sum.
[a] Computed using alpha = .05.

As shown in Table 28, mean scores of participants who indicated understanding of CCU-B, thinking like a nurse, were higher within the DML group (*M* = 34.61, *SE* = 0.680) than the non-DML group (*M* = 32.057, *SE* = 0.478). For those who received DML training, the mean DMLI sum was higher by 12.609, 95% CI[1.390, 23.827] for those who indicated understanding of the central concept represented in CCU-B than those who did not (Table 29). This was a statistically significant difference, $F(1,227) = 9.418$, $p = .028$, $n^2 = .021$.

Table 28

Mean DMLI Sum by Group Type and CCU-B (Thinking Like a Nurse) Response

Group	CCU-B Response	*N*	*M*	*SD*	*SE*	95% CI LL	95% CI UL
DML							
	No understanding	1	22.00	6.19	5.68	10.87	33.19
	Understanding	69	34.61	6.33	.684	33.26	35.96
Non-DML							
	No understanding	20	31.40	5.06	1.270	28.897	33.90
	Understanding	140	32.06	5.45	.482	31.173	33.07

Note. CI = confidence interval; LL = lower limit, UL = upper limit

Table 29

Pairwise Comparisons for Group Type and CCU-B (Thinking Like a Nurse)

Group	CCU-B		Mean Difference	*SE*	Sig.[a]	95% CI for Difference[b] LL	UL
DML							
	0	1	-12.609*	5.693	.028	-23.827	-1.390
	1	0	12.609*	5.693	.028	1.390	
	23.827						
Non-DML							
	0	1	-.657	1.351	.627	-3.320	2.005
	1	0	.657	1.351	.627	-.2.005	3.320

Note. Dependent variable: DMLI Sum. CI = confidence interval; LL = lower limit, UL = upper limit;
0 = No understanding of CCU-B,1 = Understanding of CCU-B.
* The mean difference is significant at the .05 level.
[a]Computed using alpha = .05
[b]Adjustment for multiple comparisons: Bonferroni.

CCU-C (Reflection-on-Action). A two-way ANOVA was conducted to determine if there was an interaction between the independent variables of group type and understanding of CCU-C, reflection-on-action, on DMLI sums. As shown in Table 30, the interaction effect between group type and CCU-C on DMLI sums was statistically significant, $F(1,226) = 4.538$, $p = .034$, partial $n^2 = .018$. Because of the ordinal interactions between the dependent variables, analyses of the simple main effects for CCU-C and group type were performed. A clustered bar chart (Figure 4) depicts the differences in the dependent variable (mean DMLI sums), based on the two independent variables (group type and CCU-C).

The main effects were analyzed next using Type III sums of squares to determine if an interaction effect might exist in the population. All univariate tests and pairwise comparisons were run between the cells of the simple main effects, and a Bonferroni adjustment was made to correct for multiple comparisons within each simple main effect separately. The main effect of group type on mean DMLI score was statistically significant, $F(1,226) = 12.825$, $p = .000$, partial $n^2 = .054$, and specifically with the DML group, $F(1,226) = .002$, $p = .041$, partial $n^2 = .003$. The main effect of understanding of CCU-C on the mean DMLI score was statistically significantly different, $F(1.226) = .002$, $p = .003$, partial $n^2 = .868$, specifically within the DML group, $F(1,226) = 11.420$, $p = .001$, partial $n^2 = .048$. Although the effect size was small for both main effects, both DML training and understanding of the central concept represented in CCU-C did have a statistically significant increase on the mean DMLI score.

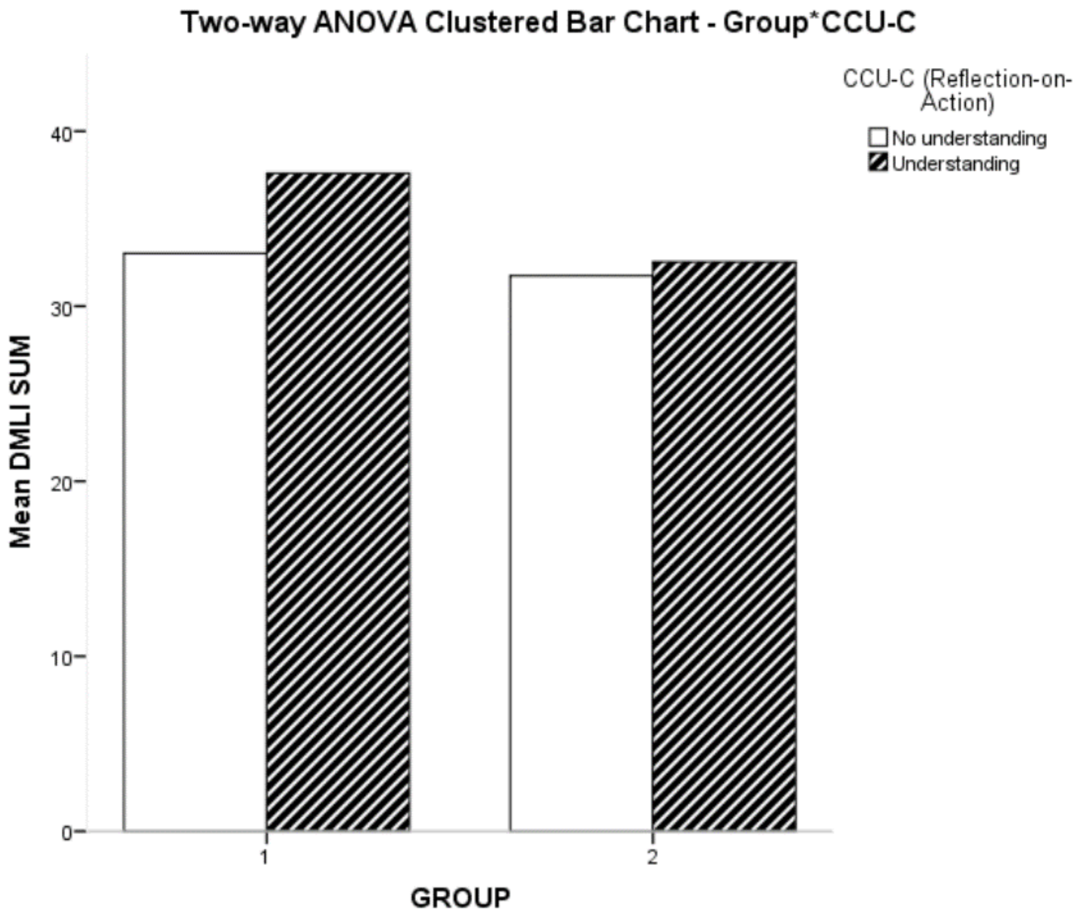


Figure 4. Two-way ANOVA (Group*CCU-C) Clustered Bar Chart with Error Bars: 95% CI. 1 = DML Group, 2 = Non-DML Group.

Table 30

Test of Between Subjects Effects for Group Type and CCU-C (Reflection-in-Action)

Source	Type III Sum Of Squares	df	Mean Square	*F*	Sig.[b]	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	591.363[a]	3	97.212	6.250	.000	.077	18.749	.963
Intercept	13452.152	1	182276.871	5776.468	.000	.651	5776.468	1.000
GROUP	43.766	1	404.687	12.825	.000	.006	12.825	.946
CCU-C	164.219	1	290.004	9.190	.003	.022	9.190	.855
GROUP*CCU-C	133.291	1	143.199	4.538	.034	.018	4.538	.564
Error	7220.778	226	31.555					
Total	253942.000	230						
Corrected Total	7678.191	229						

[a]R Squared = .077 (Adjusted R Squared = .064).
[b]Computed using alpha = .05

Mean DMLI sums of those who indicated understanding of CCU-C was highest within the DML group (M = 37.60, SE = 1.256) (Table 31). The mean score of the DML group respondents who did not indicate understanding of this central concept was higher (*M* = 33.02, *SE* = .787) than the mean score of the non-DML group who indicated understanding of the central concept (*M* = 32.53, *SE* = .819). Within the DML group, the mean score was 4.580, 95% CI[1.660, 7.501] higher for those who indicated understanding of the central concept than those who did not (Table 32), which was a statistically significant difference, $F(1,226) = 9.552$, $p = .002$, partial $n^2 = .041$ (Table 33).

Table 31

Mean DMLI Sum by Group Type and CCU-C (Reflection-on-Action) Response

Group	CCU-C Response	*N*	*M*	*SD*	*SE*	95% CI LL	95% CI UL
DML							
	No understanding	51	33.02	5.236	.787	31.47	34.57
	Understanding	20	37.60	7.810	1.256	35.13	40.08
Non-DML							
	No understanding	112	31.73	5.525	.531	30.69	32.78
	Understanding	47	32.53	5.137	.819	30.92	34.15

Note. CI = confidence interval; LL = lower limit, UL = upper limit

Table 32

Pairwise Comparisons Group Type and CCU-C (Reflection-on-Action) Response

Group	CCU-C	Mean Difference	*SE*	Sig.[a]	95% CI for Difference[b] LL	UL
DML						
0	1	-4.580*	1.482	.002	-7.501	-1.660
1	0	4.580*	1.482	.002	1.660	7.501
Non-DML						
0	1	-.800	.976	.414	-2.724	1.124
1	0	.800	.976	.414	-.1.124	2.724

Note. Dependent variable: DMLI Sum. CI = confidence interval; LL = lower limit, UL = upper limit; 0 = No understanding of CCU-C, 1 = Understanding of CCU-C.
* The mean difference is significant at the .05 level.
[a]Computed using alpha = .05
[b]Adjustment for multiple comparisons: Bonferroni.

Table 33

Univariate Tests Group Type and CCU-C (Reflection-on-Action) Response

Simple Main Effects	Sum of Squares	df	Mean Square	*F*	Sig.[a]	Partial Eta Squared	Observed Power[a]
Group Type							
Contrast	404.687	1	404.687	12.825	.000	.054	.946
Error	7131.447	226	31.555				
DML Group							
Contrast	301.403	1	301.403	.002	.041	.003	.868
Error	7131.447	226	31.555				
Non-DML Group							
Contrast	21.176	1	21.176	.671	.414	.003	.129
Error	7131.447	226	31.555				
CCU-C							
Contrast	290.004	1	290.004	9.190	.003	.039	.855
Error	7131.447	226	31.555				
No understanding							
Contrast	58.086	1	58.086	1.841	.176	.008	.272
Error	7131.447	226	31.555				
Understanding							
Contrast	360.364	1	360.364	11.420	.001	.048	.920
Error	7131.447	226	31.555				

Note. Dependent variable: DMLI Sum. [a]Computed using alpha = .05

CCU-D (Challenging Taken-for-Granted Assumptions). A two-way ANOVA was conducted to determine if there was an interaction between the independent variables of group type and understanding of CCU-D, challenging taken-for-granted assumption, on mean DMLI sums. The interaction effect between group type and CCU-D on DMLI sums was not statistically significant, $F(1,226) = 2.592$, $p = .109$, partial $n^2 = .011$ (Table 34). Although the interaction was not statistically significant, this does not mean that an interaction effect does not exist in the population. A clustered bar chart (Figure 5) depicts the differences in the dependent variable (mean DMLI sums), based on the two independent variables (group type and CCU-D).

The main effects were analyzed next using Type III sums of squares to determine if an interaction effect might exist in the population. All univariate tests (Table 35) and pairwise comparisons (Table 36) were run between the cells of the simple main effects, and a Bonferroni adjustment was made to correct for multiple comparisons within each simple main effect separately. The main effect of group type on mean DMLI sum was statistically significant, $F(1,226) = .018$, $p = .894$, partial $n^2 = .000$, specifically for the DML group $F(1,226) = 5.143$, $p = .024$, partial $n^2 = .022$. The main effect of understanding of CCU-D on the mean DMLI sum was statistically significantly different, $F(1,226) = 6.568$, $p = .011$, partial $n^2 = .028$, specifically among those who indicated understanding of CCU-D, $F(1,226) = 8.340$, $p = .004$, partial $n^2 = .036$. Although the effect size was small, understanding of the central concept represented in CCU-D did have a statistically significant increase on the mean DMLI sum despite the absence of a statistically significant interaction.

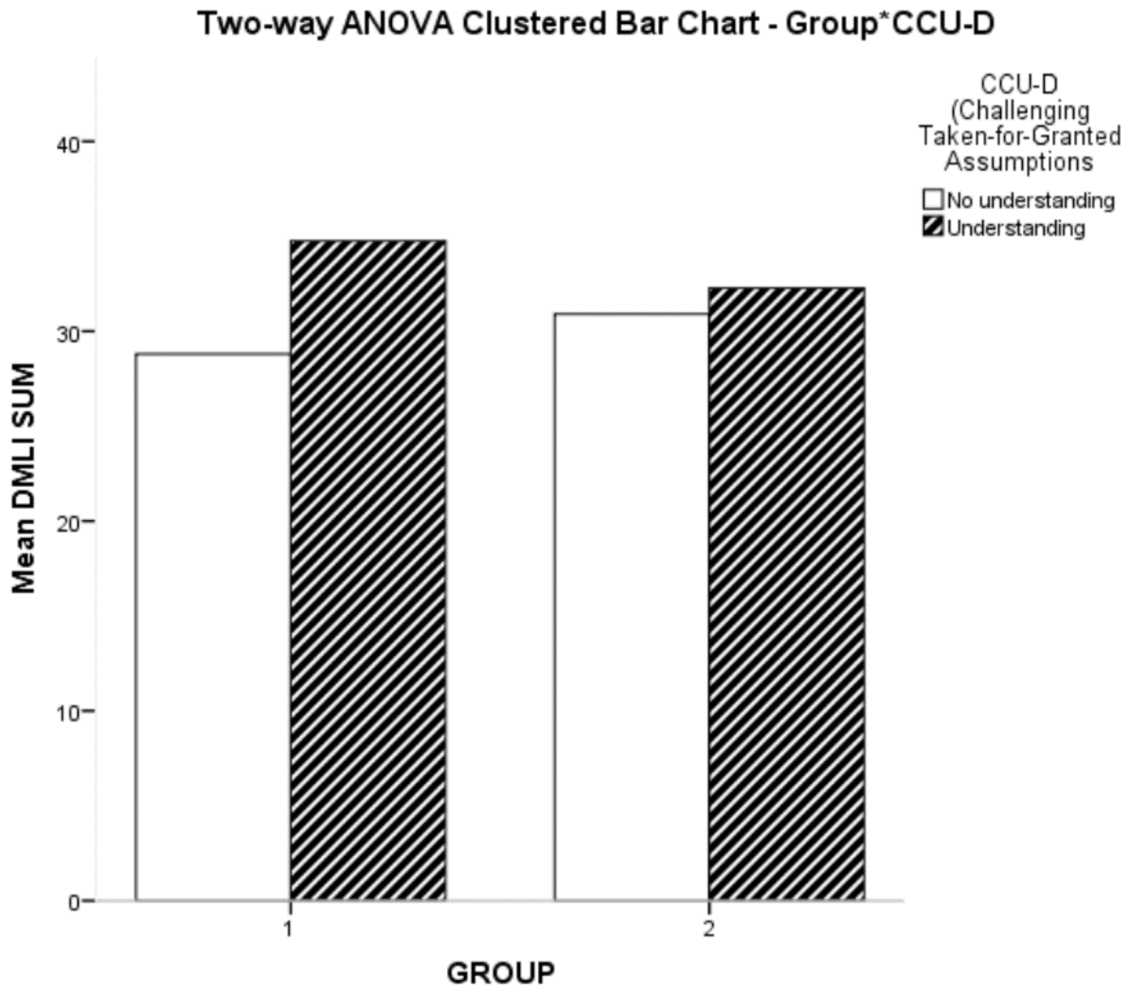


Figure 5. Two-way ANOVA (Group*CCU-C) Clustered Bar Chart with Error Bars: 95% CI. 1 = DML Group, 2 = Non-DML Group.

Table 34

Test of Between Subjects Effects for Group Type and CCU-D (Challenging Taken-for-Granted Assumptions)

Source	Type III Sum Of Squares	df	Mean Square	*F*	Sig.[b]	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	477.347[a]	3	159.116	4.972	.002	.062	14.916	.910
Intercept	63114.916	1	63114.916	1972.191	.000	.897	1972.19	1.000
GROUP	.573	1	.573	.018	.894	.000	.018	.052
CCU-D	210.194	1	210.194	6.568	.011	.028	6.568	.723
GROUP*CCU-D	82.961	1	82.961	2.592	.109	.011	2.592	.361
Error	7232.549	226	32.002					
Total	253712.000	230						
Corrected Total	7709.896	229						

[a]R Squared = .062 (Adjusted R Squared = .049).
[b]Computed using alpha = .05

Table 35

Univariate Tests for Group Type and CCU-D (Challenging Taken-for-Granted Assumptions) Response

Simple Main Effects	Sum of Squares	df	Mean Square	*F*	Sig.[a]	Partial Eta Squared	Observed Power[a]
Group Type							
Contrast	.573	1	.018	.894	.000	.018	.052
Error	7232.549	226	32.002				
DML Group							
Contrast	164.581	1	164.581	5.143	.024	.022	.617
Error	7232.549	226	32.002				
Non-DML Group							
Contrast	47.306	1	47.306	1.478	.225	.006	.228
Error	7232.549	226	32.002				
CCU-D							
Contrast	210.194	1	210.194	6.568	.011	.028	.723
Error	232.549	226	32.002				
No understanding							
Contrast	9.184	1	19.184	.599	.440	.003	.120
Error	7232.549	226	32.002				
Understanding							
Contrast	266.897	1	266.897	8.340	.004	.036	.820
Error	7232.549	226	32.002				

Note. Dependent variable: DMLI sum.
[a]Computed using alpha = .05

Mean DMLI sum of participants who responded correctly to CCU-D (challenging taken-for-granted assumptions) was higher within the DML group (*M* = 34.754, *SE* = 0.702) than the non-DML group (*M* = 32.266, *SE* = .500), as shown in Table 35. For those who indicated understanding of CCU-D within the DML group, the mean DMLI sum was higher by 5.954, 95% [.780, 11.127] than those who did not indicate understanding, which was statistically significant, $F(1,226) = 5.143$, $p = .024$, partial $n^2 = .022$ (Table 37).

Table 36

Pairwise Comparisons for Group Type and CCU-D (Challenging Taken-for-Granted Assumptions) Response

Group	CCU-D		Mean Difference	SE	Sig.[a]	95% CI for Difference[b] LL	UL
DML							
	0	1	-5.954*	2.625	.024	-11.127	-.780
	1	0	5.954*	2.625	.024	.780	11.127
Non-DML							
	0	1	-1.359	1.118	.225	-23.563	.844
	1	0	.800	.976	.414	-.844	3.563

Note. Dependent variable: DMLI Sum. CI = confidence interval; LL = lower limit, UL = upper limit; 0 = No understanding of CCU-D,1 = Understanding of CCU-D.
* The mean difference is significant at the .05 level.
[a]Computed using alpha = .05
[b]Adjustment for multiple comparisons: Bonferroni.

Table 37

Mean DMLI Sum by Group Type and CCU-D (Challenging Taken-for-Granted Assumptions) Response

Group	CCU-D Response	*N*	*M*	*SD*	*SE*	95% CI LL	UL
DML							
	No understanding	5	28.80	4.025	2.530	23.82	33.79
	Understanding	65	34.76	6.374	.702	33.37	36.14
Non-DML							
	No understanding	32	30.91	5.479	1.000	28.94	32.88
	Understanding	128	32.27	5.351	.500	31.28	33.25

Note. CI = confidence interval; LL = lower limit, UL = upper limit

CCU-E (Reflection-beyond-Action). A two-way ANOVA was conducted to determine if there was an interaction between the independent variables of group type and understanding of CCU-E (reflection-beyond-action) on DMLI scores. As presented in Table 38 the interaction effect between group type and

CCU-E on DMLI sums was not statistically significant, $F(1,226) = 2.882$, $p = .091$, partial $n^2 = .012$. The main effects were analyzed next using Type III sums of squares to determine if an interaction effect might exist in the population. A clustered bar chart (Figure 6) depicts the differences in the dependent variable (mean DMLI sums), based on the two independent variables (group type and CCU-E).

All univariate tests (Table 39) and pairwise comparisons (Table 40) were run between the cells of the simple main effects, and a Bonferroni adjustment was made to correct for multiple comparisons within each simple main effect separately. The main effect of the DML group was statistically significant, $F(1, 228) = 5.607$, $p = .019$, partial $n^2 = .024$. The main effect of not understanding CCU-E was also statistically significant, $F(1, 228) = 10.442$, $p = .001$, partial $n^2 = .044$ (Table 37).

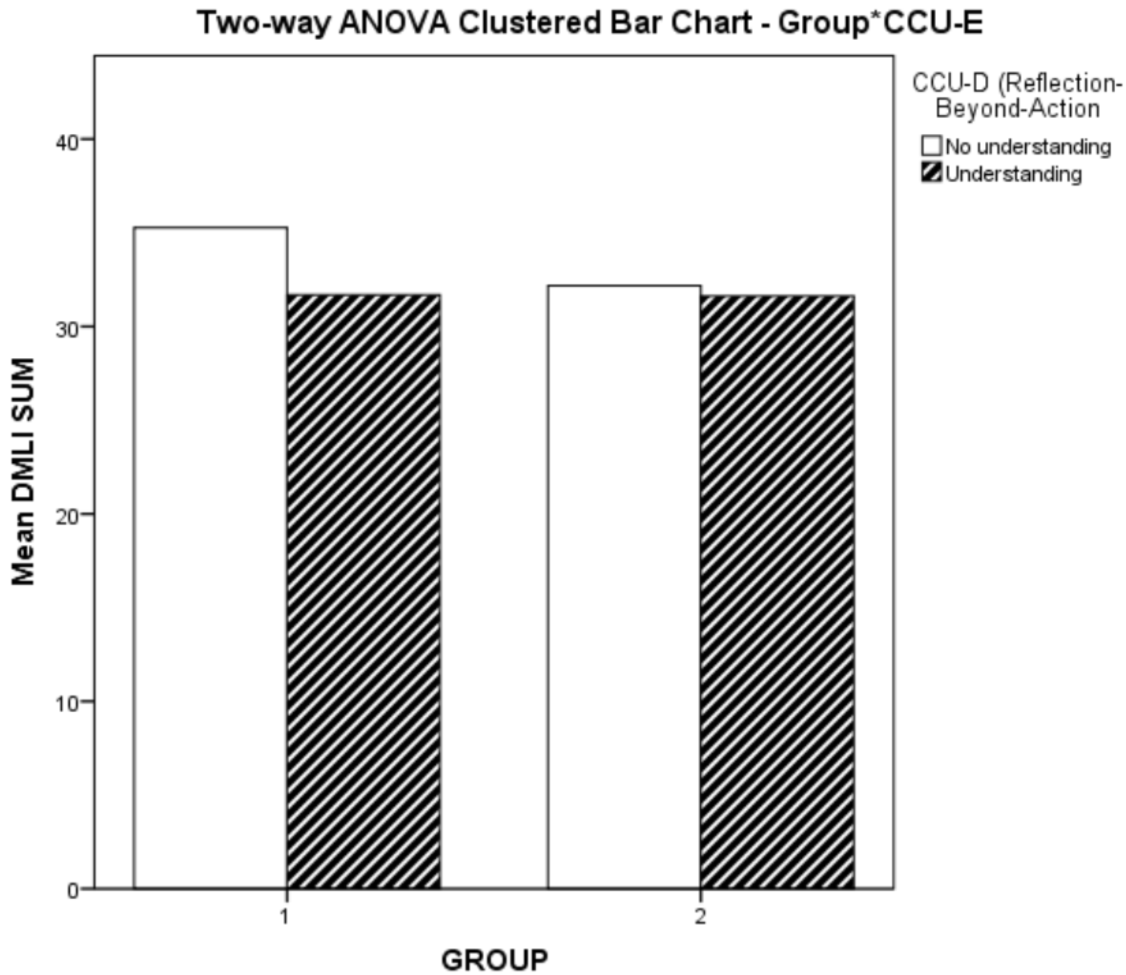


Figure 6. Two-way ANOVA (Group*CCU-E) Clustered Bar Chart with Error Bars: 95% CI. 1 = DML Group, 2 = Non-DML Group.

Table 38

Test of Between Subjects Effects for Group Type and CCU-E (Reflection-beyond-Action) Response

Source	Type III Sum Of Squares	df	Mean Square	*F*	Sig.[b]	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	455.967[a]	3	151.989	4.765	.003	.059	14.294	.897
Intercept	171854.757	1	171854.757	5387.510	.000	.959	5387.510	1.000
GROUP	100.127	1	100.127	3.139	.078	.014	3.139	.423
CCU-E	172.802	1	172.802	5.417	.021	.023	5.417	.640
GROUP*CCU-E	91.923	1	91.923	2.882	.091	.012	2.882	.394
Error	7272.912	228	31.899					
Total	255778.000	232						
Corrected Total	7728.879	231						

[a]R Squared = .059 (Adjusted R Squared = .047).
[b]Computed using alpha = .05.

Table 39

Univariate Tests for Group Type and CCU-E (Reflection-beyond-Action) Response

Simple Main Effects	Sum of Squares	df	Mean Square	*F*	Sig.[a]	Partial Eta Squared	Observed Power[a]
Group Type							
Contrast	100.127	1	100.127	3.139	.078	.014	.423
Error	7227.912	228	31.899				
DML Group							
Contrast	178.847	1	178.847	5.607	.019	.024	.655
Error	7272.912	228	31.899				
Non-DML Group							
Contrast	11.399	1	11.399	.357	.551	.002	.091
Error	7272.912	228	31.899				
CCU-E							
Contrast	172.802	1	172.802	5.417	.021	.023	.640
Error	7272.912	228	31.899				
No understanding							
Contrast	333.093	1	333.093	10.442	.001	.044	.896
Error	7272.912	228	31.899				
Understanding							
Contrast	.062	1	.062	.002	.965	.000	.050
Error	7272.912	228	31.899				

Note. Dependent variable: DMLI Sum.
[a]Computed using alpha = .05

The mean DMLI sum of participants who responded correctly to CCU-E, reflection-beyond-action, was similar between the DML group (*M* = 31.684, *SE* = 1.296) and the non-DML group (*M* = 31.618, *SE* = .762), as shown in Table 41. Within the DML group, those who indicated understanding of the central concept represented in CCU-E had a lower mean DMLI sum (*M* = 31.685, *SE* = 1.295) than those who did not indicate understanding (*M* = 35.269, *SE* = .783). The mean difference within the DML group was statistically significantly different, $F(1,226) = 5.607$, $p = .019$, partial $n^2 = .024$ (Table 40). Within the non-DML group, those who did not indicate understanding CCU-E also had a higher mean DMLI sum than those who did indicate understanding.

Table 40

Pairwise Comparisons for Group Type and CCU-E (Reflection-beyond-Action) Response

Group	CCU-E	Mean Difference	*SE*	Sig.[a]	95% CI for Difference[b] LL	UL
DML						
0	1	3.585*	1.514	.019	.602	6.568
1	0	-3.585*	1.514	.019	-6.568	-.602
Non-DML						
0	1	.561	.939	.551	-1.288	2.410
1	0	-.561	.939	.551	-2.410	1.288

Note. Dependent variable: DMLI Sum.
CI = confidence interval; LL = lower limit, UL = upper limit
0 = No understanding of CCU-E,1 = Understanding of CCU-E.
* The mean difference is significant at the .05 level.
[a]Computed using alpha = .05
[a]Adjustment for multiple comparisons: Bonferroni.

Table 41

Mean DMLI Sum by Group Type and CCU-E (Reflection-beyond-Action) Response

Group	CCU-E Response	*N*	*M*	*SD*	*SE*	95% CI LL	UL
DML							
	No understanding	52	35.27	6.362	.783	33.726	36.813
	Understanding	19	31.68	5.716	1.296	29.131	34.237
Non-DML							
	No understanding	106	32.18	5.389	.549	31.098	33.260
	Understanding	55	31.62	5.394	.762	30.118	33.119

Note. CI = confidence interval; LL = lower limit, UL = upper limit

To summarize, the findings of research question four indicate that there were statistically significant interactions between the group type, CCU-B (thinking like a nurse) and CCU-C (reflection-on-action) on the mean DMLI sums. Table 42 presents the summary of the statistically significant interactions and main effects.

Table 42

Summary of Statistical Significance of Interactions and Main Effects

CCU	Sig Main Effects	*F*	Sig.[a]	Partial Eta Squared	Observed Power[a]
CCU-A (Reflection-in-Action)					
	Group type	9.262	.003	.039	.858
	DML group	9.301	.003	.040	.859
	CCU-A	9.528	.002	.040	.867
	No understanding of CCU-A	13.138	.000	.055	.950
CCU-B (Thinking Like a Nurse)					
	Interaction Group*CCU-B	4.172	.042	.018	.529
	DML group	4.905	.028	.021	.597
	Understanding of CCU-B	5.140	.024	.022	.617
CCU-C (Reflection-on-Action)					
	Interaction Group*CCU-C	4.538	.034	.018	.564
	Group type	12.825	.000	.054	.946
	DML group	.002	.041	.003	.868
	Understanding of CCU-C	11.420	.001	.048	.920
CCU-D (Challenging Taken-for-Granted Assumptions)					
	Group type	.894	.000	.018	.052
	DML group	5.143	.024	.022	.617
	Understanding of CCU-D	8.340	.004	.036	.820
CCU-E (Reflection-beyond-Action)					
	DML group	5.607	.019	.024	.655
	CCU-E	5.417	.021	.023	.640
	No understanding of CCU-E	10.442	.001	.044	.896

[a]Computed using alpha = .05.

Summary

In this chapter, the data analysis to answer each research question was provided. A description of the DMLI and the analysis of the derived data was presented. Each of the four research questions was addressed.

Results from the first question revealed that the DMLI is a valid measure of DML understanding and application. The latent class approach to the DMLI data supported a six-class model, which provides an acceptable fit to the data. The standardized factor loadings were then examined, and R-squared was reported for each of the 52 items, in addition to the R-squared for the six DFactors that correspond to the six E-s of DML.

Findings from the second question described the debriefers' understanding of the central concepts associated with DML. There were statistically significant differences between groups for understanding of CCU-B (thinking like a nurse), and CCU-D (challenging taken for granted assumptions). The null hypothesis is rejected and the alternative hypothesis is accepted, as there were differences in DMLI sums of respondents when grouped according to the type of training received.

The third question explored the behaviors DML debriefers report consistently applying during simulation debriefing. The findings from the analysis of the DMLI data indicated that there were differences in mean DMLI sums of participants when grouped together according to the type of training receiving. The null hypothesis was rejected, and the alternative hypothesis was accepted. In addition, there were differences in mean DMLI sums in the DML group according to the source of DML training received, and according to additional training received.

The fourth question explored the interaction between the understanding of the central concepts of DML and group type, on the number of DML behaviors consistently applied during simulation debriefing with baccalaureate nursing students. There were statistically significant interactions between group type and CCU-B (thinking like a nurse), and group type and CCU-C (reflection-on-action). In addition, there was a statistically significant effect on the reported application of DML behaviors, as evidenced by mean DMLI sums, by receiving DML training. Understanding of the central concept had an effect on mean DMLI sums for CCU-B (thinking like a nurse), CCU-C (reflection-on-action), and CCU-D (challenging taken for granted assumptions), and CCU-E (reflection-beyond-action). Not understanding the central concept had an effect on mean DMLI sums for CCU-A (reflection-in-action), and CCU-E (reflection-beyond-action).

Chapter V will summarize and discuss the findings presented in this chapter. The next chapter will provide implications for nursing education. Recommendations for future research will also be addressed.

Chapter V Summary, Discussion, and Conclusions

Chapter V includes a summary of this study, a discussion of the findings, an overview of the limitations, implications for nursing education, and recommendations for further research. The intent of this chapter is to discuss the study findings, relating them to prior research in debriefing and training, and to present recommendations for further research in debriefing, within the context of nursing education.

Summary of the Study

The purpose of this study was to describe the impact of training in Debriefing for Meaningful© (DML) on how consistently trained debriefers apply and understand the central concepts of DML during simulation debriefing when grouped according to the type of training received. Training in a theory-based debriefing method has been identified as critical to debriefing facilitation (Decker et al., 2013), yet little research exists describing the training necessary for understanding and applying a debriefing method in simulation. While it is assumed that faculty who are trained to use DML debrief their students according to that training, no data exists to validate that assumption.

The Debriefing for Meaningful Learning Evaluation Scale (DMLES) was developed and tested in a pilot study as an objective measure of the debriefing behaviors consistent with the iterative process of DML. However, use of the DMLES in an adequately powered study proved too challenging to be feasible. Therefore, a subjective measure of debriefing, the Debriefing for Meaningful Learning Inventory (DMLI), based on the items of the DMLES, was developed and used to assess the impact of training in DML on how consistently trained debriefers understand and apply the central concepts of DML during simulation debriefing with baccalaureate prelicensure nursing students.

This study included four research questions to address its purpose. The first research question examined, “Is the DMLI a valid measure of DML understanding and application?” Confirmatory factor analysis (CFA) was used to determine if the DMLI is a good measure of DML. The latent class approach to analyzing the DMLI data supported a six-class model, which demonstrates an acceptable fit to the data.

The second research question asked, “Is there a difference in how many of the central concepts associated with DML that debriefers understood, when they were grouped according to the training they received?” Participant responses to these items were used to demonstrate their understanding of the central concepts of DML: reflection-in-action (CCU-A), thinking like a nurse (CCU-B), reflection-on-action (CCU-C), challenging taken-for-granted assumptions (CCU-D), and reflection-beyond-action (CCU-E). Interestingly, for each of these central concepts, there was a higher frequency of understanding within the group that was not trained to use the DML debriefing method than the group that had received training for each of the reflection central concepts. The DML group indicated understanding more frequently than the non-DML group for the concepts of thinking like a nurse, and challenging taken-for-granted assumptions.

The third question asked, “Is there a difference in how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received?” The data showed that there was evidence of a statistically significant difference in mean DMLI sum of scores between the DML group and the non-DML group.

The fourth question inquired, “Is there an interaction between the between how many of the central concepts associated with DML that debriefers understood, and how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received?” The data showed that there were statistically significant interactions between the type of training and the central concepts thinking like a nurse, and reflection-on-action, on mean DMLI sums.

Discussion of the Findings

The goal of this study was to describe the impact of training in DML on how consistently trained debriefers understand and apply the central concepts of the method. The four research questions were used to test the impact of DML training on the understanding and application of the central concepts of the method.

Research Question One

Question 1: Is the DMLI a valid measure of DML understanding and application? In order to study the impact of DML training, it was necessary to first investigate how DML is understood and applied in debriefing. While direct observation of DML debriefing behaviors using the DMLES was the original intent of this study, it was not possible to observe and rate behaviors due to lack of participation of trained debriefers. Because of the lack of debriefers who would agree to be observed and rated, the DMLI was developed and used as a self-report measure of trained debriefers' understanding and application of DML to answer the first research question. Face validity and content validity measures were used in addition to latent class factor analysis (LCFA) to determine if the DMLI is a valid measure of DML understanding and application.

Face validity of the DMLI was determined by the developer of DML and two experts in the debriefing method. Content validity was determined during testing of the DMLES, because the content had not been changed from the original 31 item DMLES from which the DMLI was derived. Items one through 52 of the DMLI represented behaviors consistent with DML; participants self-reported how consistently they applied each of these behaviors. Items 53 through 57 of the DMLI characterized the central concepts of DML; reflection-in-action, thinking like a nurse, reflection-on-action, challenging taken-for-granted assumptions, and reflection-beyond-action (Dreifuerst, 2012). Each of these five items was intended to assess whether or not a debriefer understood the central concepts of DML. These items were not included in the factor analysis because they represented broader concepts that are embedded within each of the six E's of DML.

Because it is the most common model-based clustering method used with discrete data (Dean & Raftery, 2010), LCFA was used to evaluate whether the 52 DMLI items that measured application yield good fit when grouped according to the six E's of the DML method. This was done through LCFA by examining the inter-correlations that exist between the responses to the DMLI items, then reducing the items into smaller groups called factors (Tabachnick & Fidell, 2013). The DFactor model, a type of LCFA, was used to account for the latent and discrete variables (Vermunt & Magidson, 2015).

The DMLI data supported a six-class DFactor model, yielding an acceptable fit for the data (L-squared = 7.0803, p = 0.298, BIC = 6630.79). Each of the DFactors are classified as the six E's of DML: engage, explore, explain, elaborate, evaluate, and extend. The 52 items of the DMLI that assessed application of DML behaviors loaded onto one of the six DFactors. R-squared values represent the communalities, which is the amount of variance explained in the item by the DFactor, and can be expressed as a percentage. Total R-squared values for each of the DFactors that represent the six E's of DML indicate that 78.4% of the variance is explained in the items by DFactor 1 (engage), 87.6% of the variance is explained in the items by DFactor 2 (explore), 89.8% of the variance is explained in the items by DFactor 3 (explain), 91.1% of the variance is explained in the items by DFactor 3 (elaborate), 83.1% of the variance is explained in the items by DFactor 5 (evaluate), and 76.9% of the variance is explained in the items by DFactor 6 (extend). These values reveal the 52 application items of the six E's of DML yield good fit, and indicate the DMLI is a valid measure of DML.

The findings from this first question revealed that the DMLI was a valid measure of the application of DML; this was a necessary first step prior to answering the three remaining research questions. Validity testing of the DMLI was important, because while there are debriefing tools such as the DASH© (Simon, Raemer, & Rudolph, 2010a), the DMLI was intended to measure identifiable debriefing behaviors consistent with DML rather than provide a broad measurement of the debriefing experience. Further, the discipline continues to call for more psychometrically tested instruments to further educational research (NLN, 2015).

The lack of recruitment of debriefers willing to allow an observer to behaviorally rate their DML debriefing was surprising, particularly because the International Nursing Association of Clinical Simulation and Learning (INACSL), the National League for Nursing (NLN), and the National Council of State Boards of Nursing (NCSBN) have each recommended not only training in debriefing, but also competence evaluation of debriefing. While the reasoning behind the lack of participation was not explored, this could present a challenge for future simulation education research. Without observation of a debriefer facilitating debriefing with students, it remains unknown how consistently that debriefer adheres to the original design of a specific debriefing method.

In summary, the first research question, “Is the DMLI a valid measure of DML understanding and application?” was answered first through face and content validity tests, and then through LCFA, which revealed that the 52 application items of the DMLI did yield a model of good fit.

Research Question Two

Question 2: Is there a difference in how many of the central concepts associated with DML that debriefers understood, when they were grouped according to the training they received? The findings for the second research question revealed statistically significant differences between those who were trained in DML and those who were not for two of the central concepts. Those trained in DML indicated a greater frequency of understanding than those who had not been trained in DML for two of the central concepts of DML: thinking like a nurse (p = .007), and challenging taken-for-granted assumptions (p = .015). This is important because both of these concepts are unique to DML, and are woven through each of the six E's of the method. Understanding of these two concepts is foundational to applying the iterative process of DML.

It was surprising that the non-DML group indicated a greater frequency of understanding of the three reflection central concepts. For reflection-in-action, 59.1% of the non-DML group indicated understanding compared to 49.3% of the DML group. For reflection-on-action, 29.6% of the non-DML group indicated understanding compared to 28.2% of the DML group. For reflection-beyond-action, 34.2% of the non-DML group indicated understanding compared to 26.8% of the DML group. These findings were unforeseen because it was anticipated that debriefers trained in DML would indicate a greater frequency of understanding of all of the central concepts than those not trained to use the DML method. However, with the increasing focus on reflective thinking in debriefing across simulation education (Decker et al., 2013; NCSBN, 2015), as well as across nursing curriculum (NLN, 2015), this finding may suggest that debriefers within nursing education are more aware of reflective thinking in a broad sense, regardless of the debriefing method. It may also be possible that other debriefing methods have provided exposure to the nuances of each type of reflection.

Another possible explanation for the greater frequency of the non-DML group understanding the reflection concepts of DML is that the wording of the survey items may not have been clear or consistent with the lexicon used in the DML training. In addition, debriefers may not be able to differentiate between reflection-in-action, reflection-on-action, and reflection-beyond-action, even though they indicate they use these concepts during DML debriefing. Although these are central concepts of DML, debriefers trained in the method may have learned the process of applying these types of reflection during debriefing but may not be able to identify them by name or describe their differences.

In conclusion, the answer to the second question is that there were statistically significant differences between the groups in understanding of the DML central concepts thinking-like-a nurse and challenging taken-for-granted assumptions.

Research Question Three

Question 3: Is there a difference in how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when grouped according to the training they received? The findings from the data used for the third research question indicated that there was a difference in how many DML behaviors debriefers report applying, according to whether or not they had received training in DML or in another method. Participants trained in DML engaged in more behaviors consistent with DML than participants trained in another debriefing method. This is not an unexpected finding. It would be anticipated that debriefers trained in DML would consistently apply more DML behaviors than those not trained in DML or trained in another debriefing method.

The range of DMLI sums for the DML group was 22 to 51, while the range of DMLI sums scores for the non-DML group was 17 to 42. This indicates that a debriefer trained in DML reported consistent application of DML behaviors from as little as 43% to as much as 82% of the time during debriefing with nursing students. For the DML trained debriefers, the mean DMLI sum was 34.31, indicating that as a group, only 34 of the 51 DMLI behaviors (67%) were consistently applied by DML trained debriefers. Of those not trained in DML, the mean DMLI sum was 31.90, indicating that as a group, 32 of the 51 DMLI behaviors (63%) were consistently applied by debriefers who had not received DML training.

These findings raise many questions. Educators with advanced degrees, who have been trained in a specific debriefing method, are scoring a lower than typical passing score in applying a method during debriefing, which is the time of simulation where the most significant learning should be occurring. On a typical grading scale in nursing programs, this may not be a passing grade. If debriefing training becomes further regulated by state boards of nursing, these findings suggest a need to identify a standard for an acceptable level of debriefing application for debriefers who have been trained in an evidence-based debriefing method. If simulation is indeed the most important component of simulation (Shinnick, Woo, Horwich, & Steadman, 2011), then it is important to consider how much more meaningful the learning could be among students if all debriefers were required to meet a benchmark for application of a debriefing method.

Another interesting finding for the second research question was the differences between the groups related to the sources or types of training. Again, the DML group had higher mean DMLI sums than the non-DML group for each of the reported sources, which is not unexpected. The training sources that yielded a statistically significant difference in mean DMLI sums were vendor meetings (p = .019), vendor representatives (p = .028), Center for Medical Simulation (CMS) (p = .002), and NLN Simulation Innovation Resource Center (NLN-SIRC) online courses (p = .007). Participants were not asked to provide a description of these sources of training, so the specific content and skills provided in the training participants received is unknown.

These findings are puzzling because these sources are not common sources of DML training. It is possible that much like the greater frequency of understanding of reflection concepts among the non-DML group, perhaps the training received through these various sources were heavily focused on reflection. While it is unknown who the trainers were, or the details of the training, it may be that participants found training in reflection helpful in implementing DML. It is also possible that vendor sponsored offerings could be interpreted as conferences or workshops. More data would need to be collected before further conclusions could be made.

When describing DML training they received, participants could select all applicable sources from a menu of sources within the survey to describe how they learned DML. As a single source of training, these sources included the following, which reflect lowest to highest mean DMLI sums: reading one DML article (*M* = 31.00, *SD* = 5.099), watching a colleague use DML (*M* = 34.31, *SD* = 6.360), reading more than one DML article (*M* = 35.63, SD = 6.481), attending a train-the-trainer session (*M* = 36.56, *SD* = 6.224), or attending a DML workshop or conference (*M* = 36.57, *SD* = 6.621). Participants commonly reported combining sources of training (*n* = 56). When other training sources were combined with DML workshop or conference attendance, the mean DMLI sums increased with the addition of reading more than one article (*M* = 38.00, *SD* = 6.97) and with the addition watching a colleague (*M* = 38.53, *SD* = 8.62). Furthermore, when all three of these sources were reported as sources of training (*n* = 12), the mean DMLI sums increased even more (*M* = 40.67, *SD* = 8.14). When a fourth source of training was added to this combination, train-the-trainer session attendance, (*n* = 5), the mean DMLI sum increased yet again (*M* = 41.80, *SD* = 7.19).

This is important because each of these sources as a single source of training resulted in a 60.8% to 71.7% range of application of DML behaviors, but when added to one or more additional sources of training the application of DML behaviors increased to a range of 75.5% to 82.0%. While neither the dose, timing, or repetition of training was a variable of interest in this study, these findings suggest that repeated experiences in learning DML yielded a higher application of DML behaviors. This is an important finding and is supported by reports of increased retention as a result of increasing amounts of brief deliberate practice (Oermann et al., 2011; Oermann, Kardong-Edgren, Odom-Maryon, & Roberts, 2014; Sutton, 2011; Vadnais et al., 2012), and distributed learning experiences (Raman, McLaughlin, Violato, Rostorn, Allard, and Coderre, 2010). The increasing level of DML application that corresponds with additional sources of training also suggests that knowledge retention may be linked to the varying methods of how the information was presented, as well as the dose or repetition of training in how to facilitate a DML debriefing. This finding is further supported by Ebbinghaus' (1913) seminal research on forgetting, which indicates that forgetting is impacted by how the information was learned, and how frequently the information was reviewed.

Within the DML group, it was interesting to note the similarity in the number of debriefing behaviors applied between those who received training through a DML workshop or conference (M = 36.57, SD = 6.621), and those who received training by attending a train-the-trainer session (M = 36.56, SD = 6.224). Although this was not a statistically significant difference, this indicates that those who received training through a DML workshop or conference scored 71.70% compared to a 71.69% score among those who received training through a train-the-trainer session. This finding has implications for nursing program resource allocation for training, continuing education, and faculty development. Since there was very little difference in DML application between the two sources of training, it may be more cost effective for one debriefer to attend a workshop, and then train others within the nursing program once the trainer's debriefing has been assessed to ensure consistent application of the method. The train-the-trainer approach also reinforces the learning of the debriefer who attended the conference or workshop training, through the process of teaching their new learning to others. Further study in this area is warranted.

Mean DMLI sums also demonstrated increases with the achievement of each academic agree within the DML group. Participants who reported master's preparation consistently applied more DML behaviors (M = 32.60, SD = 5.570) than baccalaureate prepared participants (M = 31.40, SD = 7.503). Those with doctoral preparation applied the highest level of DML behaviors (M = 39.32, SD = 5.386). It goes without saying that while DML training is not included in degree achievement, this finding suggests that with each successive advanced degree, a debriefer does consistently apply more of the learned DML behaviors during debriefing with baccalaureate nursing students. A possible explanation may be related to more effective cognitive encoding processes that are learned and practiced in the achievement of advanced degrees, in addition to how deeply new information is processed (Sousa, 2000).

The participants' self-perception of DML implementation was also an important finding. Mean DMLI sums rose with increasing perception of how participants believed they implemented DML. None of the participants *strongly disagreed* that they implemented DML well, and only few (n = 9) *disagreed* (M =

31.56, *SD* = 5.270). In fact, as self-perception of DML application increased, the mean DMLI sums also increased, as demonstrated by the participants who indicated they *agreed* (*n* = 45) that they implement DML well (*M* = 33.93, *SD* = 6.770), and participants who *strongly agreed* (*n* = 14) that they implement DML well (*M* = 37.57, *SD* = 4.636). Although self-report can be limiting in nature due to self-perception (Paulhus & Vazire, 2007), the increase in mean DMLI sums suggests that the participants' indication of perception of how they implement DML, was consistent with how their DMLI sums demonstrated they apply DML.

Within the DML group, the source of training that demonstrated the highest DMLI mean sum was found among the group trained for the NCSBN National Simulation Study (NSS) (Hayden et al., 2014). Although only four debriefers from the NSS trained group participated in this study, their mean DMLI sum was the highest of the DML single sources of training (*M* = 38.75, *SD* = 2.630). The training of this group was described by Jeffries and colleagues (2015) as including intensive instruction with repeated DML learning experiences. This group was also evaluated for DML competence numerous times throughout the two years of the study, with remediation as necessary. It is important to note that the NSS study ended two years prior to the time these participants responded to the DMLI, indicating there was not a substantial loss of knowledge and skills retention as may have been expected. It is possible that their intense dispersed learning experiences with regular competency assessment promoted a deeper processing of the application of DML, which was reflected in their DMLI sums.

These findings are important because as regulations and guidelines are developing for simulation and debriefing, the data begin to demonstrate that a one-time offering of debriefing training may not lead to consistently accurate use of the debriefing method in the manner that demonstrated evidence of improved student learning outcomes. In addition, re-assessment of trained debriefers using an instrument that has demonstrated a statistically significant good fit to the method is a critical piece of guidelines and regulations, which thus far has been missing in the recommendations.

In conclusion, the findings of the third research question derived from the DMLI data indicate that there are differences in the application of DML behaviors based on the types and sources of training debriefers received.

Research Question Four

Question 4: Is there an interaction between how many of the central concepts associated with DML that debriefers understood, and how many behaviors associated with DML debriefers report they consistently apply during simulation debriefing, when they were grouped according to the training they received? The fourth research question examines the interaction between understanding of each of the five central concepts of DML, and the type of training received on the application of DML. There were statistically significant interactions between DML training and understanding of the central concepts thinking like a nurse (p = .042) and reflection-on-action (p = .034), on the application of DML. For each of the central concepts, the main effects were also examined to understand how the variables interacted. Interaction effects represent the combined effects of the group type (DML or non-DML), and each central concept on application of DML (mean DMLI sums).

Reflection-in-Action. The mean DMLI difference was statistically significant (p = .003) between the groups, indicating that although the interaction of group type and understanding of the concept did not have a statistically significant interaction on how consistently debriefers apply DML (p = .051), there was a statistically significant effect by DML training (p = .003), and the lack of understanding of the concept (p = .000), on mean DMLI sums. This means that in spite of DML training, not understanding the concept of reflection-in-action impacts the application of DML behaviors.

Thinking Like a Nurse. There was a statistically significant interaction between group type and understanding of thinking like a nurse on how consistently debriefers apply DML (p = .042). DML training had a statistically significant effect on application of DML (p = .028), as did understanding of the concept (p = .002). There was a statistically significant difference in mean DMLI sums between groups for those who indicated understanding of this concept (p = .028). In other words, those who were trained in DML and indicated understanding the concept of thinking like a nurse, consistently applied a higher number of DML behaviors.

It is an important finding that there was an interaction between understanding thinking like a nurse, and having received training in the DML debriefing method. As a central concept of DML, thinking like a nurse embodies clinical reasoning, and describes the complex and non-linear process of thinking and reasoning a nursing student must learn in preparation for the complexity of patient care that is evident in nursing practice (Dreifuerst, 2012; Tanner, 2006). This finding suggests that debriefers' understanding of this concept had an impact on their application of DML when debriefing nursing students. Because other debriefing methods do not emphasize the concept of thinking like a nurse, this differentiates DML as specific to changing thinking and reasoning processes within the domain of nursing. Learning to think like a nurse in the context of clinical care is challenging and this method of debriefing may help students achieve this (Dreifuerst, 2012; Forneris et al., 2015; Mariani, Cantrell, Meakim, Prieto, & Dreifuerst, 2013).

Reflection-on-Action. There was a statistically significant interaction between group type and understanding of the concept reflection-on-action on DML application (p = .034). The group type had a statistically significant effect on the application of DML (p = .000), specifically if debriefers were trained in DML (p = .041), and if there was understanding of the concept (p = .001). There was also a statistically significant difference (p = .002) in the mean DMLI sums between groups for those who indicated understanding of this concept. Therefore, those who were trained in DML and indicated understanding of reflection-on-action consistently applied a higher number of DML behaviors when debriefing nursing students. Guiding a learner to understand how to reflect on thoughts and actions, and to analyze the connection between the two, is important not only in DML application, but also represents the development of a reflective practitioner (Schön, 1983).

Challenging Taken-for-Granted Assumptions. Although there was no statistically significant interaction between group type and the concept of challenging taken-for-granted assumptions on the application of DML, the group type had a statistically significant effect on the application of DML (p = .000), specifically if debriefers were trained in DML (p = .024). Understanding of the concept of challenging taken-for-granted assumptions also had a statistically significant effect on the application of DML by (p = .004). There was a statistically significant difference (p = .024) in the mean DMLI sums between groups for those who indicated understanding of this concept. In other words, those who were trained in DML and indicated understanding of the concept challenging taken-for-granted assumptions consistently apply a higher number of DML behaviors when debriefing nursing students.

Reflection-beyond-Action. Although there was not a statistically significant interaction between group type and the central concept of reflection-beyond-action, the main effect of DML training had a statistically significant effect on DML application (p = .019). Not understanding the concept of reflection-beyond-action also had a statistically significant effect on the application of DML (p = .001). There was a statistically significant difference in the mean DMLI sums between groups for those who did not indicate understanding of this concept. In spite of receiving training in DML, not understanding the central concept of reflection-beyond-action impacted the application of DML behaviors.

These findings are important for the future of training in DML. Even if debriefers receive training, regardless of the number of times, not understanding the central concepts of the method can impact how DML is applied during debriefing with students. Although the non-DML group indicated understanding of the concepts more frequently than the DML group, training in DML still resulted in applying more DML behaviors.

In summary, the findings of the fourth research question indicate that there were statistically significant interactions between the type of debriefing training (DML or non-DML) and understanding the central concepts of the method, which impacted the application of DML when debriefing prelicensure baccalaureate nursing students.

Implications for Nursing Education

The findings from this study have several implications for nursing education and nurse educators. The NCSBN issued guidelines stating that "simulation education in nursing programs should be based on educational theories" (Alexander et al., 2015, p. 41). DML is one of the few debriefing methods that meet this criterion. Furthermore, the *INACSL Standards of Best Practice: Simulation*^{SM} stated in Standard VI that a debriefer should receive formal training in a structured debriefing method, and that competence should be validated through established instruments (Decker et al., 2013). The NLN (2015) also recommended formal training in a theory-based debriefing method, with on-going assessment of competence for all debriefers, to promote debriefing across the nursing curriculum. Clearly, there is agreement that debriefers need training in rigorous methods. Undoubtedly, there is also agreement that competence of debriefing skills should be evaluated. The findings of this study, however, present substantial challenges to the implementation, although not the intent, of these recommendations.

If debriefers are trained in a theory-based debriefing method, it cannot be assumed that the training is robust. There are many ways that debriefers can be trained to use evidence-based debriefing methods, yet this study demonstrated that training does not necessarily equate with consistent application of the debriefing method in the manner intended. Even if the training is robust, this does not guarantee that the debriefing method is being applied as it was intended. In this study, debriefers who had been trained to use DML still applied the method in a variety of ways that were not necessarily consistent with the originally tested design of DML. Therefore, if debriefing training is required, application of the method must also be evaluated. However, the findings of this study suggest the need to determine a minimum required standard of debriefing application that must be achieved by trained debriefers. This is necessary because, for instance, DML is one debriefing method that has demonstrated a positive impact on student learning outcomes (Dreifuerst, 2012; Forneris et al., 2015; Mariani et al., 2013). Nevertheless, if DML is not applied explicitly as it was designed, it is unknown what the impact will be on student learning outcomes. It may not be enough to merely require training in a theory-based debriefing method; determining a minimum standard of debriefing application is a critical next step, with a subsequent plan for re-training to achieve that standard.

It is important to note that the DML group was asked to describe their experiences regarding evaluation of their DML debriefing. When asked if their DML training included an evaluation with feedback, 35% reported they had been evaluated (*n* = 24) while 65% reported they had not (*n* = 45). Participants were also asked to describe how many times their DML debriefing had been evaluated since receiving training. Forty-two percent reported they had never been evaluated, 32% reported evaluation one to three evaluations of application, 7% reported evaluation four to six evaluations, and 10% reported evaluation greater than six times. It is unknown what these evaluations consisted of, how the DML debriefing was measured, nor if there were implications regarding the outcome. These reports of assessment indicate that some debriefers are evaluated, in alignment with the recommendations of the regulating bodies, but many are not. Nevertheless, the mean DMLI sums reported in this study are a result of the trained debriefers who actually have had their application of DML evaluated (*n* = 35/71) since receiving training. This suggests that perhaps merely evaluating after training may not be enough.

Requiring a minimum standard is not new to the discipline of nursing. Within nursing education, nursing students graduate with the intent of passing National Council Licensure Examination (NCLEX-RN®) at a minimum passing standard. However, this minimum passing standard reflects "nursing ability required to practice competently at the entry level" (NCSBN, 2016). Similarly, it has yet to be acknowledged, much less determined, how to define a minimum passing standard that indicates an acceptable level of understanding and application of a debriefing method. This study revealed that the mean DMLI sum of the DML trained debriefers was 34.31 (67%) which would not be an acceptable score for nursing students on any assessment. This highlights the need for requiring more and improved training and evaluation, with a standard for debriefing application.

Another implication from the study findings is the impact of the additive effect of multiple sources of training on the mean DMLI sums. Based on the data, it is clear that debriefers benefit from more than a 'one-and-done' approach to training in evidence-based debriefing methods in order to meet the spirit of the recommendations from INACSL, the NLN, and NCSBN. While this study did not investigate nor reveal the necessary dose and repetition interval necessary, it did point to the importance of determining this and including a recommendation in future debriefing guidelines.

Finally, another implication of this study is a need for deep contemplation by the academe as well as regulation bodies, regarding the assumption that faculty who are debriefers will apply what they learn during debriefing training when they debrief students. This study challenges this assumption. As faculty, the aspect of academic freedom needs to be considered, particularly in the context of the increased focus on building an evidence-based teaching practice across higher education, but specifically in nursing education (Ferguson & Day, 2005; Ironside & Valiga, 2008; Kalb et al., 2015; McAllister, Oprescu, & Jones, 2014; Patterson, 2009). A tension exists within nursing education between adhering to evidence-based teaching practices and integrating individual academic freedom. As nurse clinicians, a nurse is required to adhere to all steps of an evidence-based nursing intervention in practice with patients. For example, in clinical practice a nurse would not be evaluated positively if there was only a 67% application of an evidence-based sterile technique, yet the findings of this study indicate that this occurs commonly during debriefing of nursing students who are being prepared for a high-stakes healthcare profession. This raises questions regarding the lack of utilization and assessment of evidence-based teaching practices in the preparation of students for an evidence-based health profession.

Many participants in this study included a written description of their debriefing practice which indicated that debriefers commonly combine or adapt methods, modify DML, or add their own self-determined debriefing techniques. In fact, the debriefing strategy, Promoting Excellence and Reflective Learning in Simulation (PEARLS), advocates for a blended approach to debriefing and encourages educators to purposefully merge various debriefing strategies to tailor discussion to learner needs and learning context (Eppich & Cheng, 2015). However, there is no evidence to demonstrate that an individually blended or debriefer-created approach yields the same consistent learning outcomes as using a tested debriefing method, thus defying the purpose of basing debriefing on evidence-based teaching practices. Again, if academic freedom and faculty choice takes prominence over adhering to the evidence, it is unlikely that any individual combination of methods will yield consistent learning outcomes without testing these specific combinations. This too needs further exploration and extensive research.

Limitations

Several limitations were identified within this study. The first is related to the sample. This convenience sample was derived from an open recruitment of members of INACSL who debrief prelicensure baccalaureate nursing students, in addition to other debriefers known to the investigator. This sample may not be representative of all debriefers who debrief in nursing education. Debriefers who were likely to have received debriefing training were recruited, but it is possible that there are other debriefers who have received training but are not members of INACSL who may have provided additional data regarding types and sources of debriefing training.

The use of a self-reported survey was another limitation. The DMLI data derived from self-report could have challenges related to validity and reliability, as responses could have been influenced by emotion, bias, and interpretation (Paulhus & Vazire, 2007).   In the original study design, debriefers would have been observed and scored as they debriefed prelicensure baccalaureate nursing students. Lack of recruitment of participants willing to have their debriefing observed and rated, or administrators willing to have their faculty debriefing observed and rated, made that original observational design unfeasible. Debriefing, by nature, is a confidential environment between the debriefer and learners. Although some programs record debriefing and some may implement peer review, there are security and confidentiality measures in place to protect the privacy of all involved. Perhaps because of the private context of debriefing, debriefers were reluctant to be observed and rated. In addition, there may be fear of repercussion by faculty and administrators if the debriefer is scored low. Nonetheless, without observation of the debriefer, how a debriefer applies a debriefing method will remain unknown. While use of the DMLI, a self-reported version of the objective DMLES, provided useful data, an objective measure might have yielded different results.

A third limitation was related to clarity of the types and sources of training. There may have been confusion regarding how participants interpreted and differentiated between a conference or workshop, training that occurred at vendor meetings, or trainings at specific locations that were named as sources. Confusion between the sources of training may have skewed the results if participants did not consistently identify a conference or workshop similarly.

Recommendations for Further Research

The purpose of this study was to investigate the impact of training in DML on the understanding and application of the method in simulation debriefing. Future research is needed to further explore all aspects of debriefing training. First, the measurement of debriefing behaviors could be explored further with the DMLI and the DMLES. The DMLI could be normalized to provide understanding of the score that should be expected from a debriefer who consistently applies the method. Also, direct observation of a DML debriefing could be measured by an observer with the DMLES, then compared to the debriefer’s self-reported DMLI measure. This would provide understanding of how similarly the DMLES and the DMLI measure DML debriefing behaviors, as well as comparing a debriefers self-perception of application to an observed measure of application. This study was conducted with debriefers of baccalaureate nursing students; differences in debriefers of other populations of nursing students could provide additional information in the understanding and application of DML after training. Further testing of the DMLES and the DMLI in various contexts would provide additional validity of these instruments.

Another recommendation for future research is assessing pre-knowledge prior to DML training with the DMLI. The findings of the non-DML group suggest that perhaps debriefers who have not been trained in DML possess knowledge of the central concepts, and are applying DML behaviors without having received training in the method. Evaluating pre-knowledge with the DMLI would give insight into a debriefer's self-report of how they describe their debriefing behaviors. Comparison of DMLI scores pre and post training would provide understanding of any differences produced through training.

Another suggestion for additional research involves investigating other variables that could impact a trained debriefer's application of a debriefing method. The dispersed effect of training, dose, and re-dose in the context of debriefing are all unexplored. Research is needed to determine when best to train, and to re-train, debriefers to promote the highest level of retention. Testing the dose and re-dose of each of the sources of training that yielded higher mean DMLI sums would provide insight into differences between types of sources, to accommodate all types of learners.

This study made evident that debriefing behaviors consistent with DML can be measured through objective behavioral measurement, as determined through the previous pilot testing of the DMLES, and through subjective measurement, as demonstrated with the DMLI. The ability to measure debriefing behaviors, as opposed to limiting measurement to broad debriefing constructs, provides a mechanism for evaluation, has implications for training, and could identify debriefing behaviors that promote learning outcomes.

Additional research should also consider development of measurement tools for other evidence-based, reflective debriefing methods. Because participants reported receiving training in other debriefing methods, the application of these methods could be evaluated with developed and tested tools. This could provide data that would describe commonalities between the application of different debriefing methods, to guide the future of required debriefing characteristics. More instruments are needed within the discipline that demonstrate valid and reliable measurement of debriefing.

Secondary analysis of the DMLI data could also provide useful information in the future. Further examination of the qualitative data collected is warranted to describe more fully the combinations of debriefing types, sources and timing of training, and other training characteristics. This information may provide greater insight into all variables that impact the understanding and application of DML in debriefing.

Finally, there were many comments within the survey from participants regarding the use of various components of multiple debriefing methods during a debriefing. These comments need further exploration and from an analysis, a study could be designed to explore the impact on student learning from this style of debriefing practice.

Conclusion

The findings from this research study contribute to the work of previous researchers in the area of debriefing in nursing education by expanding on prior work to include methods of assessing debriefers' debriefing practices. This study also deepened understanding of how debriefers are receiving training to facilitate debriefing in nursing education. The findings revealed the impact of common types and sources of both DML and non-DML training on the debriefer's understanding and application of central components of DML. This investigation was the first study to describe how debriefers apply an evidence-based debriefing method in debriefing, specifically DML, and to provide a valid instrument for measurement. DML has been adopted widely across nursing education, but it has been unknown how the method has been translated into debriefing practice.

In conclusion, this study uncovered valuable information regarding the impact of debriefer training on debriefing behaviors. It contributes to the growing body of knowledge of debriefing in nursing education, and training in evidence-based debriefing methods.

Appendix A Institutional Review Board Approval

INDIANA UNIVERSITY
OFFICE OF THE VICE PRESIDENT FOR RESEARCH
Office of Research Compliance

To: Kristina Dreifuerst
NURSING

Cynthia Bradley
UNIVERSITY LEVEL

From:

Human Subjects Office
Office of Research Compliance – Indiana University

Date: November 23, 2015

RE: **NOTICE OF EXEMPTION - NEW PROTOCOL**

Protocol Title:	Exploring DML Debriefing with the Debriefing for Meaningful Learning Evaluation Scale (DMLES) in Prelicensure Nursing Education
Study #:	1511761795
Funding Agency/Sponsor:	None
Status:	Exemption Granted | Exempt

Study Approval Date: November 23, 2015

The Indiana University Institutional Review Board (IRB) EXE000001 | Exempt recently reviewed the above-referenced protocol. In compliance with (as applicable) 45 CFR 46.109 (d) and IU Standard Operating Procedures (SOPs) for Research Involving Human Subjects, this letter serves as written notification of the IRB's determination.

Under 45 CFR 46.101(b) and the SOPs, as applicable, the study is accepted as Exempt (1) Category 1: Educational Research Conducted in Educational Settings. 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, with the following determinations:

Acceptance of this study is based on your agreement to abide by the policies and procedures of the Indiana University Human Research Protection Program and does not replace any other approvals that may be required. Relevant policies and procedures governing Human Subjects Research can be found at: http://researchcompliance.iu.edu/hso/hs_guidance.html.

The Exempt determination is valid indefinitely. Substantive changes to approved exempt research must be requested and approved prior to their initiation. Investigators may request proposed changes by submitting an amendment through the KC IRB system. The changes are reviewed to ensure that they do not affect the exempt status of the research. Please check with the Human Subjects Office to determine if any additional review may be needed.

You should retain a copy of this letter and all associated approved study documents for your records. Please refer to the assigned study number and exact study title in future correspondence with our office. Additional information is available on our website at http://researchcompliance.iu.edu/hso/index.html.

Appendix B Faculty Study Information Sheet

IRB STUDY # 151761795

INDIANA UNIVERSITY STUDY INFORMATION SHEET FOR FACULTY
Exploring Debriefing in Prelicensure Nursing Education

STUDY PURPOSE
You are invited to participate in a research study designed to explore the different ways nurse educators facilitate simulation debriefing with prelicensure nursing students. Please read this form and ask any questions you may have before agreeing to participate in the study.

This study is being conducted by Cynthia Sherraden Bradley PhD(c), RN, CNE, CHSE, a doctoral candidate at Indiana University School of Nursing, under the supervision of Kristina Thomas Dreifuerst PhD, RN, CNE, ANEF, an Assistant Professor at Indiana University School of Nursing.

PROCEDURES FOR THE STUDY
If you agree to be in this study, you will be asked to **<u>complete an electronic survey</u>** describing your experience in simulation debriefing with prelicensure nursing students. First, you will be asked to answer demographic questions describing your preparation for, and experience with simulation debriefing. No personally identifying information will be asked of you, to maintain complete anonymity. Next, you will be asked a series of questions regarding how you facilitate a simulation debriefing session.

Organizations that may inspect and/or copy your research records for quality assurance and data analysis include groups such as the study investigator and his/her research associates, and the Indiana University Institutional Review Board or its designees, the study sponsor, and (as allowed by law) state or federal agencies, specifically the Office for Human Research Protections (OHRP).

PAYMENT
You will not receive payment for taking part in this study. We greatly appreciate your willingness to participate in this study that we believe will make a contribution to the understanding of debriefing.
BENEFITS OF TAKING PART IN THE STUDY
There is no direct benefit to you for participating in this study.

CONTACTS FOR QUESTIONS OR PROBLEMS
For questions about the study, contact the researchers Cynthia Sherraden Bradley MSN, RN, CNE, CHSE at [redacted] or [redacted] or Kristina Thomas Dreifuerst PhD, RN, CNE, ANEF at [redacted] or [redacted].
For questions about your rights as a research participant or to discuss problems, complaints or concerns about a research study, or to obtain information, or offer input, contact the IU Human Subjects Office at (317) 278-3458 or (800) 696-2949.

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

SUBJECT'S CONSENT

Since consent is not required in IRB Exempt studies, your agreement to participate in the study according to these terms is presumed when you click on link below to access the electronic survey.

Appendix C Debriefing for Meaningful Learning Inventory (DMLI)

Options: *Always, Sometimes, Never*

1. During debriefing with my students, I allow quiet students to silently reflect and not engage in the discussion.
2. While debriefing with my students, I spend time correcting misconstruction of information or action.
3. During debriefing with my students, I focus on remaining nonjudgmental.
4. During debriefing, I engage my students to discuss the patient's story including what they learned in report and what they uncovered during the patient encounter.
5. While debriefing my students, I use a Socratic style of questioning.
6. During debriefing with my students, I offer suggestions of what the learner may have been thinking during decision-making.
7. While debriefing with my students, I write notes from the discussion on a white board or something similar.
8. During debriefing with my students, I focus the discussion on the objectives of the simulation.
9. While debriefing with my students, I ask them to describe how the care of the patient in the simulation applies to other patients.
10. While debriefing with my students, I provide mini-lectures based on the scenario content.
11. During debriefing with my students, I share with students what I might have been thinking had I been in the same patient situation.
12. While debriefing with my students, I aim to keep all students engaged.
13. During debriefing with my students, I provide an evaluation of each student's performance during the simulation.
14. During debriefing with my students, I ask questions that test the learner's preparedness for this simulation.
15. During debriefing, I encourage my students to identify clinical decisions made during the simulation.
16. While debriefing with my students, I ask them what could have been done differently in the simulation.
17. When debriefing, I am focused on uncovering their thinking of my students regardless whether the actions I observed were correct or incorrect.
18. While my students are debriefing, I encourage them to debrief themselves.
19. While debriefing with my students, I use questions that were prepared in advance of the simulation experience.
20. While debriefing my students, I guide them to reflectively think back on the assumptions that impacted the clinical decisions and actions they made.
21. While debriefing with my students, I ask questions that prompt them to discuss their understanding of the patient's disease process (ie – pathophysiology, pharmacology).
22. While debriefing with my students, I encourage discussion by pairing advocacy with inquiry.

23. During debriefing with my students, simulation observers are expected to engage in the discussion as actively as simulation participants.
24. During debriefing, I ask my students progressive questions based on their responses to focus our discussion on exploring the students' thinking.
25. During debriefing, I have my students write on worksheets.
26. While debriefing with my students, I encourage them to explain the decision-making processes they used.
27. During debriefing, I encourage students to reflect on their thoughts that led to a decision.
28. During debriefing, I intentionally give each student feedback on their performance in simulation.
29. While debriefing with my students, I encourage them to reflect on the assessments performed during the simulation.
30. During debriefing with my students, I encourage them to describe what they were thinking that prompted an action I observed during the simulation.
31. While debriefing with my students, I use the term thinking like a nurse.
32. While debriefing with my students, I encourage them to point out the mistakes of other participants.
33. During debriefing with students, we discuss the care of other types of patients similar to the patient in the simulation scenario.
34. While debriefing with my students, I summarize the patient's story for all of us.
35. During debriefing with my students, I encourage them to explain why they chose not to do an anticipated a nursing action.
36. After debriefing with my students, I ask them to write in a journal about the simulation experience.
37. While debriefing with my students, I will interrupt them, particularly when they are wrong or the discussion is going in a direction other than what I intend.
38. During debriefing, I encourage my students to reflect on their thoughts, decisions, and actions during the simulation.
39. During debriefing, I encourage my students to extend their thinking beyond this scenario by presenting a parallel scenario of a similar patient situation.
40. While debriefing with my students, I spend the majority of debriefing discussing the incorrect clinical decisions and actions made during the simulation (what went wrong).
41. During debriefing with my students, I provide a lecture if I feel it is necessary.
42. While debriefing with my students, I encourage them to verbalize what went right and what wrong during the simulation.
43. During debriefing, I encourage my students to think about what they will do differently the next time they encounter this situation.
44. While debriefing with my students, I encourage them to identify critical decision-making points in the simulation
45. While debriefing with my students, I encourage them to reflectively think about, and identify the factors that contributed to decisions and actions during the simulation.
46. During debriefing with my students, I ask them to explain their thinking associated with actions and decisions I observed in the simulation.

47. While debriefing with my students, I encourage them to reflect on the patient's response to the decisions actions made during the simulation.
48. During debriefing with my students, I use a technique to help them park their emotions.
49. During debriefing with my students, I ask them to identify the nursing skills they need to practice for future patient care experiences.
50. While debriefing with my students, I am focused on uncovering the thinking behind their actions during the simulation.
51. When debriefing with my students, I am consistent in my approach in order to develop their thinking routines.
52. During debriefing with my students, I emphasize anticipating, assimilating, and accommodating.

Options: *Yes*, *No*

53. When I am debriefing with my students, reflection-in-action is when we talk about how the care of this patient applies to the care of another patient.
54. When I am debriefing with my students, when we talk about thinking like a nurse this means how nurses use knowledge and experience to approach a patient condition.
55. When I am debriefing with my students, when we talk about reflection on action we are talking about those clinical decision points that happen in the moment.
56. When I am debriefing with my students, when I am challenging taken for granted assumptions I am interested in matching thinking with actions.
57. When I am debriefing with my students and I create an alternative patient scenario with an unrelated diagnosis, I am asking them to think beyond action.

Appendix D DMLES

Debriefing for Meaningful Learning Evaluation Scale	Present 1	Not Present 2
The facilitator engaged participants by:		
Gathering to begin debriefing		
Guiding participants thru beginning the DML worksheet		
Providing 2-3 minutes to silently journal initial thoughts		
Drawing in all learners throughout the debriefing session		
Redirecting conversation throughout debriefing session to keep learners engaged		
Maintaining a listening posture throughout debriefing session		
Abstaining from a lecture format throughout debriefing session		
Framing the patient story		
The facilitator guided participants in exploring by:		
Identifying the clinical decisions made		
Identifying contributing factors to making decisions		
Identifying responses to the clinical decisions		
Guiding thinking about correct actions		
Identifying what could have been done differently		
Guiding thinking about incorrect actions		
Using Socratic questioning to uncover thinking		
Challenging taken-for-granted assumptions		
The facilitator guided participants in explaining by highlighting:		
Assessments performed		
Decision-making processes used		
Actioned decisions made		
Patient responses to decisions made		

Making relevant connections		
Connecting care of this patient to a broader scope of these types of patients		
The facilitator encouraged participants in elaborating by:		
Expanding thinking and actions		
Engaging in reflection-in-action		
Engaging in reflection-on-action		
The facilitator collaborated with participants in evaluating by:		
Verbalizing what went right/wrong actions		
Identifying critical points of the experience		
The facilitator extended thinking by:		
Engaging participants in reflection-beyond-action		
Asking 'what-if' questions		
Creating an alternate scenario		
Linking common concepts from current scenario to parallel scenario		

Appendix E Demographic Questions

Type of teaching institution

- University/college
- Hospital-based nursing program

Highest academic degree completed

- Bachelor
- Master
- Doctoral

Years teaching traditional BSN students

- Less than 1
- 1-5
- 6-10
- 11-15
- 16-20
- 21 or more

Years using debriefing in simulation

- Less than 1
- 1-2
- 3-4
- 5-6
- 7-8
- 9-10
- 10 or more

Type of training received for debriefing:

- Mentor
- Conference or workshop
- Vendor representative
- Vendor meeting
- Local/Regional training
- NLN-SIRC online courses
- INACSL Webinar
- WISER center
- SSH-CHSE program
- Center for Medical Simulation (Harvard)

- Center for Health Science Inter-Professional Education, Research and Practice (University of Washington)
- Graduate-level certificate program
- Michael S. Gordon for Research in Medical Education
- Other
- None

Since receiving debriefing training, my debriefing has been evaluated:

- Never
- 0-3 times
- 4-6 times
- 6 or more times

The debriefing method I use is:

- Debriefing for Meaningful Learning
- Debriefing with Good Judgment
- Gather-Analyze-Summarize (GAS)
- National League for Nursing (NLN) 3 phase process
- 3D model (Defusing, Discovering, & Deepening)
- Outcome-Present-State Test (OPT) model
- Other (please specify)

For respondents who report using DML:

I use DML in simulation debriefing with prelicensure nursing students:

- 2-4 times per semester
- 5-8 times per semester
- >8 times per semester

I learned to use DML through: (Check all that apply)

- DML workshop or conference (If YES, then go to description of training)
- Read one article
- Read more than one article
- Watched a colleague use DML in debriefing
- Attended a train-the-trainer session (please describe):
- Other (please describe):

The DML workshop I attended was offered by (organization):

The DML workshop was led by (name of presenter)

Month/Year of DML training:

Length of training session:

Training session included my demonstration of DML for evaluation or feedback: *yes* *no*

Since receiving training, my implementation of DML has been evaluated:

- Never
- 0-3 times
- 3-6 times
- >6 times

I believe that I implement DML well:

- Strongly agree
- Agree
- Disagree
- Strongly Disagree

References

Alexander, M., Durham, C. F., Hooper, J. I., Jeffries, P. R., Goldman, N., Kardong-Edgren, S., Kesten, K. S., Spector, N., Tagliareni, E., Radtke, B., & Tillman, C. (2015). NCSBN simulation guidelines for prelicensure nursing programs. *Journal of Nursing Regulation, 6*(3), 39-42. doi: http://dx.doi.org/10.1016/S2155-8256(15)30783-3.

Anitei, M., & Chraif, M. (2011). Using behaviorally anchored rating scales as criteria for validating the Undergraduate Students Stress Inventory for the young Romanian Psychology students. *International Proceedings of Economics Development & Research, 23.* doi:0.1016/j.sbspro.2011.11.412

Arora, S., Ahmed, M., Paige, J., Nestel, D., Runnacles, J., Hull, L., & Sevdalis, N. (2012). Objective structured assessment of debriefing: bringing science to the art of debriefing in surgery. *Annals of Surgery, 256*(6), 982-988. doi: 10.1097/SLA.0b013e3182610c9.

Astin, A. W. (1968). Undergraduate achievement and institutional "excellence". *Science, 161*(842), 661-668.

Astin, A. W. (1993). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education.* Phoenix: The Oryx Press.

Astin, A. W. (2012). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education*. Rowman & Littlefield Publishers.

Astin, A. W., & Sax, L. J. (1998). How undergraduates are affected by service participation. *Higher Education, 39*(3), 251-263.

Banning, M. (2008). Clinical reasoning and its application to nursing: Concepts and research studies. *Nurse Education in Practice, 8*(3), 177-183. doi: 10.1016/j.nepr.2007.06.004

Bartels, J. (2007). Preparing nursing faculty for baccalaureate-level and graduate-level nursing programs: Role preparation for the academy. *Journal of Nursing Education, 46*(4), 154-158.

Benner, P. (1984). *From Novice to Expert*. Menlo Park, CA: Addison-Wesley.

Benner, P., Sutphen, M., Leonard, V., & Day, L. (2010). *Educating nurses: A call for radical transformation,15*. John Wiley & Sons.

Berndt, J., Dinndorf-Hogenson, G., Herheim, R., Hoover, C., Lang, N., Neuwirth, J., & Tollefson, B. (2015). Collaborative classroom simulation (CCS): An innovative pedagogy using simulation in nursing education. *Nursing Education Perspectives, 36*(6), 401-402. doi: 10.5480/14-1420

Billings, D. M., & Halstead, J. A. (2013). *Teaching in nursing: A guide for faculty.* Elsevier Health Sciences.

Billings, C. E, Reynard, W. D. (1984). Human factors in aircraft incidents: results of a 7-year study. *Aviation Space Environmental Medicine*, *55*, 960–965.

Bok, D. (2009). *Our underachieving colleges: A candid look at how much students learn and why they should be learning more*. Princeton University Press.

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research, 17*(3), 303-316. doi: 10.1177/0049124189017003004

Boud, D., Keogh, R., & Walker, D. (1985). *Reflection: Turning experience into learning*. Routledge Falmer.

Boyd, E. M., & Fales, A. W. (1983). Reflective learning key to learning from experience. *Journal of Humanistic Psychology, 23*(2), 99-117. doi: 10.1177/0022167883232011

Bradley, C. S., & Dreifuerst, K. T. (2016). Pilot testing the debriefing for meaningful learning evaluation scale. *Clinical Simulation in Nursing, 12*(7), 277-280. doi: 10.1016/j.ecns.2016.01.008

Brett-Fleegler, M., Rudolph, J., Eppich, W., Monuteaux, M., Fleegler, E., Cheng, A., & Simon, R. (2012). Debriefing assessment for simulation in healthcare: Development and psychometric properties. *Simulation in Healthcare, 7*(5), 288-294. doi: 10.1097/SIH.0b013e3182620228.

Bruning, R., Schraw, G., & Norby, M. (2010). *Cognitive psychology and instruction.* Columbus.

Bybee, R. W. (1989). *Science and Technology Education for the Elementary Years: Frameworks for Curriculum and Instruction*.

Bybee, R. W., Taylor, J. A., Gardner, A., Van Scotter, P., Powell, J. C., Westbrook, A., & Landes, N. (2006). *The BSCS 5E instructional model: Origins and effectiveness*. Colorado Springs, CO: BSCS, 5, 88-98.

Campbell, J., & Blakely, L. S. (1996). Assessing the impact of early remediation in the persistence and performance of underprepared college students. In *Association for Institutional Research Forum, Albuquerque, NM*.

Cangelosi, P. R., Crocker, S., Sorrell, J. M. (2009). Expert to novice: Clinicians learning new roles as clinical nurse educators. *Nursing Education Perspectives, 30*(6), 367-371.

Cantrell, M. A. (2008). The importance of debriefing in clinical simulations. *Clinical Simulation in Nursing, 4*(2), e19-e23. doi: 10.1016/j.ecns.2008.06.006

Cappell, C. L., & Kamens, D. H. (2002). Curriculum assessment: A case study in sociology. *Teaching Sociology*, 467-494.

Cheng, A., Rodgers, D. L., van der Jagt, É., Eppich, W., & O'Donnell, J. (2012). Evolution of the Pediatric Advanced Life Support course: enhanced learning with a new debriefing tool and Web-based module for Pediatric Advanced Life Support instructors. *Pediatric Critical Care Medicine, 13*(5), 589-595. doi: 10.1097/PCC.0b013e3182417709

Conigliaro, R. L., & Stratton, T. D. (2010). Assessing the quality of clinical teaching: a preliminary study. *Medical Education, 44*(4), 379-386. doi: 10.1111/j.1365-2923.2009.03612.x

Davis, D. A., Mazmanian, P. E., Fordis, M., Van Harrison, R.R., Thorpe, K.E., & Perrier L. (2006). Accuracy of physician self-assessment compared with observed measures of competence: A systematic review. *Journal of the American Medical Association, 296*(9), 1094–1102.

doi :10.1001/jama.296.9.1094

De Regge, M., Calle, P., De Paepe, P. & Monsieurs, K. (2008). Basic life support refresher training of nurses: Individual training and group training are equally effective. *Resuscitation, 79*, 283-287. doi: 10.1016/j.resuscitation.2008.06.015

Dean, N., & Raftery, A. E. (2010). Latent class analysis variable selection. *Annals of the Institute of Statistical Mathematics, 62*(1), 11-35. doi: 10.1007/s10463-009-0258-9.

DeChamplain, A. F. (2010). A primer on classical test theory and item response theory for assessments in medical education. *Medical Education, 44*(1), 109–117. doi:10.1111/j.1365-2923.2009.03425.x

Decker, S., Fey, M., Sideras, S., Caballero, S., Rockstraw, L. (R.), Boese, T., Franklin, A. E., Gloe, D., Lioce, L., Sando, C. R., Meakim, C., & Borum, J. C. (2013). Standards of Best Practice: Simulation Standard VI: The debriefing process. *Clinical Simulation in Nursing, 9*(6S), S27-S29. http://dx.doi.org/10.1016/j.ecns.2013.04.008.

Dent, J., & Harden, R. M. (2013). *A practical guide for medical teachers*. Elsevier Health Sciences.

DeVellis, R. F. (2003). *Scale development: Theory and applications*. Thousand Oaks, CA: Sage.

Dewey, J. (1910). The American Association for the Advancement of Science. *Science, 31*(787), 121-127.

Dismukes, R. K., Gaba, D. M., & Howard, S. K. (2006). So many roads: facilitated debriefing in healthcare. *Simulation in Healthcare, 1*(1), 23-25.

Dowdy, E., Twyford, J., Sharkey, J. D. (2013). Methods of assessing behavior: Observations and rating scales. T*he Oxford Handbook of Child Psychological Assessment*. doi: 10.1093/oxfordhb/9780199796304.013.0026

Dreifuerst, K. T. (2009). The essentials of debriefing in simulation learning: a concept analysis. *Nursing Education Perspectives, 30*(2), 109-114.

Dreifuerst, K. T. (2010). *Debriefing for meaningful learning: Fostering development of Clinical reasoning through simulation*. (Order No. 3617512, Indiana University). ProQuest Dissertations and Theses, 212. Retrieved from http://search.proquest.com/docview/1527174151?accountid¼7398 (1527174151).

Dreifuerst, K. T. (2012). Using debriefing for meaningful learning to foster development of clinical reasoning in simulation. *The Journal of Nursing Education, 51*(6), 326. doi: 10.3928/01484834-20120409-02

Dreifuerst, K. T. (2015). Getting started with Debriefing for Meaningful Learning. *Clinical Simulation in Nursing, 11*(5), 268-275. doi: 10.1016/j.ecns.2015.01.005

Dreifuerst, K.T. & Bradley, C.S. (2015). Transformation of teaching and learning through Debriefing for Meaningful Learning. In G. Sherwood & S. Horton-Deutsch, (Eds.) *Reflective organizations: On the Frontlines to Transform*

Education and Practice. Indianapolis, IN: Sigma Theta Tau International.

Dreifuerst, K.T., & Decker, S. (2012). Debriefing: An essential component for learning in simulation pedagogy. In P.R. Jeffries (Ed.) *Simulation in nursing education: From conceptualization to evaluation (2nd Edition*), New York: The National League for Nursing

Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology*. Ruger HA, Bussenius CE, translator. New York: Teachers College, Columbia University.

Eckert, J. (2014). Teach like a novice: Lessons from beginning teachers. *Phi Delta Kappan 2014*(96), 13-18. doi: 10.1177/0031721714553404

Elder, L., & Paul, R. (1998). The role of Socratic questioning in thinking, teaching, and learning. *The Clearing House, 71*(5), 297-301.

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*,*10*(2), 106-115. doi: 10.1097/SIH.0000000000000072

Fanning, R. M., & Gaba, D. M. (2007). The role of debriefing in simulation-based learning. *Simulation in Healthcare, 2*(2), 115-125. doi: 10.1097/SIH.0b013e3180315539

Faraway, J. J. (2015). *Linear models with R (2nd ed.).* Boca Raton, FL: CRC Press.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Ferguson, L. F., & Day, R. A (2005). Evidence-based nursing education: Myth or reality? *Journal of Nursing Education, 44*(3), 107-115.

Fey, M. K. (2014). *Debriefing practices in nursing education programs in the United States* (Doctoral dissertation). Retrieved from ProQuest (3621880).

Fincher, R. M. E., & Work, J. A. (2006). Perspectives on the scholarship of teaching. *Medical Education-Oxford, 40*(4), 293-295.

Fitzmaurice, M. (2010). Considering teaching in higher education as a practice. *Teaching in Higher Education, 15*(1), 45-55. doi: 10.1080/13562510903487941

Forneris, S. G., Neal, D. O., Tiffany, J., Kuehn, M. B., Meyer, H. M., Blazovich, L. M., Smerillo, M. (2015). Enhancing clinical reasoning through simulation debriefing: A multisite study. *Nursing Education Perspectives, 36*(5), 304-310.

Fox, J. (2008). *Applied regression analysis and generalized linear models (2nd ed.).* Thousand Oaks, CA: Sage Publications, Inc.

Garden, A. L., LeFevre, D. M., Waddington, H. L., Weller, J. M. (2015). Debriefing after simulation-based non-technical skill training in healthcare: A systematic review of effective practice. *Anaesthesia & Intensive Care, 43*(3)

Gardner, S.S. (2014). From learning to teach to teaching effectiveness: Nurse educators describe their experiences. *Nursing Education Perspectives, 35*(2), 106-111.

Groccia, J.E., Buskist, W. (2011). Need for evidence-based teaching. *New Directions for Teaching and Learning, 128*, 5-11. doi: 10.1002/tl.463

Gurung, R. A., & Schwartz, B. M. (2009). Pedagogical research: Focusing on learning. *Optimizing Teaching and Learning: Practicing Pedagogical Research*, 87-144.

Gururaja, R. P., Yang, T., Paige, J. T. & Chauvin, S. W. (2008). Examining the effectiveness of debriefing at the point of care in simulation-based operating room team training. In Henriksen, K., Battles, J.B., Keyes, M.A. & Grady, M.L. (Eds.), *Advances in Patient Safety: New Directions and Alternative Approaches, 3*,1–18.

Harden, R. M., & Crosby, J. R. (2000). The good teacher is more than a lecturer: the twelve roles of the teacher. *Medical Teacher, 22*(4), 334-347.

Hawkins, R., Bendickson, L., Benson, P., Osborne, L., McPherson, J., Todd, L., Snelson, J., Bruner, S., Bohan, K. (2014). A pilot study evaluating the perceptions of Certified Registered Nurse Anesthetists toward human patient simulation. *American Association of Nurse Anesthetists, 82*(5), 375-384.

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 nursing education. *Journal of Nursing Regulation, 5*(2), S1-S64.

House, J. D. (1999). The effect of entering characteristics and instructional experiences on student satisfaction and degree completion and degree completion: An application. *International Journal of Instructional Media, 26*(4), 423.

Howell, D. C. (2010). *Statistical methods for psychology (7th ed., international edition).* Belmont, CA: Wadsworth, Cengage Learning.

Ironside, P. M., & Valiga, T. M. (2008). National survey on excellence in nursing education. *Nursing Education Perspectives, 27*(3), 166-169.

Ja'afar, R. (2012). Two decades of championing faculty development: Is it worth the effort? *Education in Medicine Journal, 4(*2), e1-e6. doi:10.5959/eimj.v4i2.17

Jeffries, P. R. (2005). A framework for designing, implementing and evaluating simulations used to teach strategies in nursing. *Nursing Education Perspectives, 26*, 96-103.

Jeffries, P. R. (2016). *The NLN Jeffries Simulation Theory*. Wolters Kluwer.

Jeffries, P. R., & Rodgers, K. J. (2007). *Theoretical framework for simulation design. Simulation in nursing education: From conceptualization to evaluation*, 21-33.

Jeffries, P. R., Dreifuerst, K. T., Kardong-Edgren, S., & Hayden, J. (2015). Faculty development when initiating simulation programs: Lessons learned from the National Simulation Study. *Journal of Nursing Regulation, 5*(4), 17-23.

Kalb, K.A. (2008). Core competencies of nurse educators: Inspiring excellence in nurse educator practice. *Nursing Education Perspectives, 29*(4), 217-219.

Kalb, K.A., O'Conner-Von, S.K., Brockway, C., Pierson, C.L., Sendelbach, S. (2015). Evidence-based teaching practice in nursing education: Faculty perspectives and practices. *Nursing Education Perspectives, 36*(4), 212-215. doi: 10.5480/14-1472

Kautz, D. D., Kuiper, R., Pesut, D. J., Knight-Brown, P., & Daneker, D. (2005). Promoting clinical reasoning in undergraduate nursing students: application and evaluation of the Outcome Present State Test (OPT) model of clinical reasoning. *International Journal of Nursing Education Scholarship, 2*(1).

Kelly, L. J. (1996). *Implementing Astin's IEO model in the study of student retention: a multivariate time dependent approach*. United States Coast Guard Academy, Center for Advanced Studies.

Kerfoot, B. P., & Brotschi, E. (2009). Online spaced education to teach urology to medical students: a multi-institutional randomized trial. *American Journal of Surgery*, *197*(1), 89-95. doi:10.1016/j.amjsurg.2007.10.026

Klein, S. (1986). Socratic dialogue in the Meno. *Southern Journal of Philosophy, 24*, 351-363.

Kolb, D. (1984). *Experiential Learning Experience as the Source of Learning and*

Development. Prentice Hall, Englewood Cliffs: New Jersey

Kottner, J., Audigé, L., Brorson, S., Donner, A., Gajewski, B. J., Hróbjartsson, A., Streiner, D. L. (2011). Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *Journal of Clinical Epidemiology, 64*(1), 96–106.

Kuiper, R. A., & Pesut, D. (2004). Promoting cognitive and metacognitive reflective reasoning skills in nursing practice: self-regulated learning theory. *Journal of Advanced Nursing, 45*(4), 381. doi: 10.1046/j.1365-2648.2003.02921.x

Kuiper, R., Pesut, D., & Kautz, D. (2009). Promoting the self-regulation of clinical reasoning skills in nursing students. *The Open Nursing Journal, 9*, 76-85. doi: 10.1046/j.1365-2648.2003.02921.x.

Lai, N. M., & Teng, C. L. (2011). Self-perceived competence correlates poorly with objectively measured competence in evidence based medicine among medical students. *BioMed Central Medical Education, 11*, 25. doi:10.1186/1472-6920-11-25.

Lapkin, S., Levett-Jones, T., Bellchambers, H., & Fernandez, R. (2010). Effectiveness of patient simulation manikins in teaching clinical reasoning skills to undergraduate nursing students: A systematic review. *Clinical Simulation in Nursing, 6*(6), e207-e222.

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

Lazarsfeld, P. F., Henry, N. W., & Anderson, T. W. (1968). *Latent structure analysis*. Boston: Houghton Mifflin.

Lederman, L. C. (1992). Debriefing: Toward a systematic assessment of theory and practice. *Simulation & Gaming, 23*(2), 145-160.

Levett-Jones, T., Hoffman, K., Dempsey, J., Jeong, S. Y. S., Noble, D., Norton, C. A. & Hickey, N. (2010). The 'five rights' of clinical reasoning: An educational model to enhance nursing students' ability to identify and manage clinically 'at risk' patients. *Nurse Education Today, 30*(6), 515-520.

Lucas, N., Macaskill, P., Irwig, L., Moran, R., & Bogduk, N. (2009). Reliability of physical examination for diagnosis of myofascial trigger points: A systematic review of the literature. *The Clinical Journal of Pain, 25*(1), 80-89.

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382-386.

Magidson, J., & Vermunt, J. K. (2004). Latent class models. *The Sage handbook of quantitative methodology for the social sciences*, 175-198.

Malott, K. M., Hall, K. H., Sheely-Moore, A., Krell, M. M., Cardaciotto, L. (2014). Evidence-based teaching in higher education: Application to counselor education. *Counselor Education and Supervision, 53*, 294-297. doi: 10.1002/j.1556-6978.2014.00064.x

Mariani, B., Cantrell, M. A., Meakim, C., Prieto, P., & Dreifuerst, K. T. (2013). Structured debriefing and students' clinical judgment abilities in simulation. *Clinical Simulation in Nursing, 9*(5), e147-e155.

Mariani, B., Cantrell, M.A., Meakim, C. (2014). Nurse educators' perceptions about structured debriefing in clinical simulation. *Nursing Education Perspective, 35*(5), 330-331. doi: 10.5480/13-1190.1

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective, 1.* Psychology Press.

McAllister, M., Oprescu, F., Jones, C. (2014). N2E: Envisioning a process to support transition from nurse to educator. *Contemporary Nurse 46*(2), 242-250.

McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., & Scalese, R. J. (2010). A critical review of simulation-based medical education research: 2003–2009. *BioMed Central Medical Education, 44*(1), 50-63.

Meakim, C., Boese, T., Decker, S., Franklin, A. E., Gloe, D., Lioce, L., Sando, C. R., & Borum, J. C. (2013). *Standards of Best Practice: Simulation Standard I: Terminology. Clinical Simulation in Nursing, 9*(6S), S3-S11. doi: http://dx.doi.org/10.1016/j.ecns.2013.04.001.

Mezirow, J. (1981). A critical theory of adult learning and education. *Adult Education Quarterly, 32*(1), 3-24.

McKnight, P. E., McKnight, K. M., Sidani, S., & Figueredo, A. J. (2007). *Missing data: A gentle introduction.* Guilford Press.

Michell, J. (2000). Normal science, pathological science and psychometrics. *Theory & Psychology,10*, 639–667.

Monk, D.H. (2015). Reflections on teacher preparation. *Sociology, 52*, 219-224. doi: 10.1007/s12115-015-9889-z

Murphy. J. I. (2004). Using focused reflection and articulation to promote clinical reasoning: An evidence-based teaching strategy. *Nursing and Health Care Perspectives, 25*, 226–231.

National Council of State Boards of Nursing. (2016). *Passing Standard*. Retrieved from https://www.ncsbn.org/2630.htm

National League for Nursing Board of Governors. (2015). *Debriefing across the curriculum: A living document from the National League for Nursing in collaboration with the International Nursing Association for Clinical Simulation and Learning*.

Neill, M. A., & Wotton, K. (2011). High-fidelity simulation debriefing in nursing education: A literature review. *Clinical Simulation in Nursing, 7*(5), e161-e168.

Nguyen, T. H., Han, H.-R., Kim, M. T., & Chan, K. S. (2014). An introduction to item response theory for patient-reported outcome measurement. *The Patient, 7*(1), 23–35.

Nori, M. J., Saghafinia, M., Motamedi, M. H. K., & Hosseini, S. M. K. (2012). CPR training for nurses: How often is it necessary? *Iranian Red Crescent Medical Journal, 74*(2), 104-107

Oermann, M. H., Kardong-Edgren, S. E., Odom-Maryon, T, Hallmark, B. F., Hurd, D., Rogers, N., Haus, C., McColgan, J. K., Snelson, C., Dowdy, S. W., Resureccion, L.A., Kuerschner, D. R., LaMar, J., Tennant, M. N., Smart, D. A. (2011). Deliberate practice of motor skills in nursing education: CPR as exemplar. *Nursing Education Perspectives*, *32*(5), 311-315. doi:10.5480/1536-5026-32.5.311

Oermann, M. H., Kardong-Edgren, S. E., Odom-Maryon, T., Roberts, C. J. (2014). Effects of practice on competency in single-rescuer cardiopulmonary resuscitation. *Medsurg Nursing, 23*(1), 22-28.

Oleson, A., Hora, & M. T. (2014). Teaching the way they were taught? Revisiting the sources of teaching knowledge and the role of prior experience in shaping faculty teaching practices. *Higher Education (2014),68*, 29-45. doi: 10.1007/s10734- 013-9678-9

O'Steen, D. S., Kee, C. C., Minick, H. P: (1996). The retention of advanced cardiac life support knowledge among registered nurses. *Journal of Nursing Staff Division,12*, 6–72.

Overholser, J. C. (1993). Elements of the Socratic method: I. Systematic questioning. *Psychotherapy: Theory, Research, Practice, Training, 30*(1), 67.

Paige, J. T., Arora, S., Fernandez, G., & Seymour, N. (2015). Debriefing 101: Training faculty to promote learning in simulation-based training. *The American Journal of Surgery, 209*(1), 126-131.

Patterson, B.J. (2009). The nature of evidence in teaching practice. *Journal of Nursing Education, 48*(6), 327.

Paulhus, D. L., & Vazire, S. (2007). The self-report method. *Handbook of Research Methods in Personality Psychology, 1*, 224-239.

Pearson M, & Smith D. (1986). Debriefing in experience-based learning. *Simulation/Games for Learning 1986, 16*, 155–172.

Pesut, D. J., & Herman, J. (1999). OPT: Transformation of nursing process for contemporary practice. *Nursing Outlook, 46*(1), 29-36.

Pike, G. R., & Killian, T. S. (2001). Reported gains in student learning: Do academic disciplines make a difference? *Research in Higher Education, 42*(4), 429-454.

Postareff, L., Lindblom-Ylänne, S., & Nevgi, A. (2007). The effect of pedagogical training on teaching in higher education. *Teaching and Teacher Education, 23*(5), 557-571.

Raemer, D., Anderson, M., Cheng, A., Fanning, R., Nadkarni, V., & Savoldelli, G. (2011). Research regarding debriefing as part of the learning process. *Simulation in Healthcare, 6*(7), S52-S57.

Raju, D., Su, X., & Patrician, P. A. (2014). Using Item Response Theory models to evaluate the practice environment scale. *Journal of Nursing Measurement, 22*(2), 323-341.

Raman, M., McLaughlin, K., Violato, C., Rostom, A., Allard, J. P., & Coderre, S. (2010). Teaching in small portions dispersed over time enhances long-term knowledge retention. *Medical Teacher, 32*(3), 250-255.

Reed, S. J. (2012). Debriefing experience scale: Development of a tool to evaluate the student learning experience in debriefing. *Clinical Simulation in Nursing, 8*(6), e211-e217. doi: 10.1016/j.ecns.2011.11.002

Ricahrdson, H., Goldsamt, L. A., Simmons, J., Gilmartin, M., & Jeffries, P. R. (2014). Increasing faculty capacity: Findings from an evaluation of simulation clinical teaching. *Nursing Education Perspectives, 35*(5), 308-314. doi:10.5480/14-1384

Ritchhart, R., Church, M., & Morrison, K. (2011). *Making thinking visible: How to promote engagement, understanding, and independence for all learners*. John Wiley & Sons.

Rudolph, J. W., Simon, R., Raemer, D. B., & Eppich, W. J. (2008). Debriefing as formative assessment: Closing performance gaps in medical education. *Academic Emergency Medicine, 15*(11), 1010-1016.

Sankar, J., Vijayakanthi, N., Sankar, M. J., & Dubey, N., (2013), Knowledge and skill retention or in-service versus preservice nursing professionals following an informal training program in pédiatrie cardiopulmonary resuscitation: A repeated-measures quasi-experimental study. *BioMed Research international. Online publication*. doi:10.1155/2013/ 403415

Sawatzky, J. A. V., & Ennis, C. L. (2009). A mentoring needs assessment: Validating mentorship in nursing education. *Journal of Professional Nursing, 25*(3), 145-150.

Scanlan, J. M., & Chernomas, W. M. (1997). Developing the reflective teacher. *Journal of Advanced Nursing, 25*(6), 1138-1143.

Schmid, W. T. (1983). Socratic moderation and self-knowledge. *Journal of the History of Philosophy, 21*, 339-348.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action (Vol.5126)*. Basic books.

Searle, S. R. (2006). *Linear models for unbalanced data*. Hoboken, NJ: John Wiley & Sons, Inc.

Shinnick, M. A., Woo, M., Horwich, T. B., & Steadman, R. (2011). Debriefing: The most important component in simulation? *Clinical Simulation in Nursing, 7*(3), e105-e111.

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

Shulman, L. S. (2005). Signature pedagogies in the professions. *Daedalus, 134*(3), 52-59.

Simmons, B. (2010). Clinical reasoning: Concept analysis. *Journal of Advanced Nursing, 66*(5), 1151-1158. doi: 10.1111/j.1365-2648.2010.05262.x

Simmons, B., Lanuza, D., Fonteyn, M., Hicks, F., & Holm, K. (2003). Clinical reasoning in experienced nurses. *Western Journal of Nursing Research, 25*, 701-719. doi: 10.1177/0193945903253092

Simon, R., Raemer, D.B., Rudolph, J. W. (2010a) *Debriefing Assessment for Simulation in Healthcare© – rater version*. Boston, MA Center for Medical Simulation.

Simon, R., Raemer, D. B., Rudolph, J. W. (2010b). *Debriefing Assessment for Simulation in Healthcare (DASH)© – Student Version, Long Form*. Center for Medical Simulation, Boston, Massachusetts.

Simon, R., Raemer, D. B., Rudolph, J. W. (2010c). *Debriefing Assessment for Simulation in Healthcare (DASH)© – Instructor Version, Long Form*. Center for Medical Simulation, Boston, Massachusetts.

Simon R., Raemer D. B., Rudolph J. W. (2010d). *Debriefing Assessment for Simulation in Healthcare (DASH)© – Student Version, Short Form*. Center for Medical Simulation, Boston, Massachusetts.

Simon R., Raemer D. B., Rudolph J. W. (2010e). *Debriefing Assessment for Simulation in Healthcare (DASH)© – Instructor Version, Short Form*. Center for Medical Simulation, Boston, Massachusetts.

Smith, K. K., Gilcreast, D, & Pierce, K. (2008). Evaluation of staff's retention of ACLS and BLS skills. *Resuscitation, 78*, 59-6.

Sousa, D. A. (2011). *How the Brain Learns*. Corwin Press.

Sutton, R. M., Niles, D., Meaney, P. A., Aplenc, R., French, B., Abella, B. S., Lengetti, E. L., Berg, R. A., Helfaer, M. A., Nadkarni, V. (2011). Low-dose, high-frequency CPR training improves skill retention of in-hospital pediatric providers. *Pediatrics, 128*(1), e145-e151.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics (6th ed.*). Pearson: Upper Saddle River, NJ.

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

Thorndike, R. M. & Thorndike-Christ, T. (2010). *Measurement and evaluation in psychology and education (8th ed.)*. Boston, MA: Prentice Hall.

Thurmond, V. A., & Popkess-Vawter, S. (2001). Examination of a middle range theory: Applying Astin's input-environment-outcome (IEO) model to web-based education. *Online Journal of Nursing Informatics, 7*(2).

Vadnais, M. A., Dodge, L. E., Awtrey, C. S., Ricciotti, H. A., Golen, T. H., & Hacker, M. R. (2012). Assessment of long-term knowledge retention following single-day simulation training for uncommon but critical obstetrical events. *Journal of Maternal-Fetal & Neonatal Medicine, 25*(9), 1640-1645. doi:10.3109/14767058.2011.648971

Vermunt, J. K., & Magidson, J. (2005). *Latent GOLD 4.0 User's Guide*. Belmont, Massachusetts: Statistical Innovations Inc.

Vermunt, J. K., & Magidson, J. (2015). U*pgrade Manual for Latent GOLD 5.1* Belmont, Massachusetts: Statistical Innovations Inc.

Walker, L. O., & Avant, K. C. (2005). *Strategies for theory construction in nursing.* Upper Saddle River, NJ: Pearson.

Walker, L. O., Avant, K. C. (2011). *Strategies for Theory Construction in Nursing.* Upper Saddle River, NJ: Pearson.

Warne, R., T., McKyer, E. L. J., & Smith, M., L. (2012). An introduction to item response theory for health behavior researchers. *American Journal of Health Behavior, 36*(1), 31–43.

Wayne, D. B., Moazed, F., Cohen, E. R., Sharma, R. K., McGaghie, W. C., & Szmuilowicz, E. (2012). Code status discussion skill retention in internal medicine residents: One-year follow-up. *Journal of Palliative Medicine*, *15*(12), 1325-1328. doi:10.1089/jpm.2012.0232

Waznonis, A. R. (2014). Methods and evaluations for simulation debriefing in nursing education. *Journal of Nursing Education, 53*(8), 459-465.

Waznonis, A. R. (2015). Simulation debriefing practices in traditional baccalaureate nursing programs: National survey results. *Clinical Simulation in Nursing, 11*(2), 110-110. doi: 10.1016/j.ecns.2014.10.002.

Zell, E., & Krizan, Z. (2014). Do people have insight in their abilities? A metasynthesis. *Perspectives on Psychological Sciences, 9*(2), 111-125.

Curriculum Vitae

Cynthia Diane Bradley

EDUCATION

2016	Ph.D., Indiana University
2013	M.S. Nursing, University of Central Missouri
1991	B.S. Nursing, University of Kansas

LICENSURE

Registered Nurse, Missouri

PROFESSIONAL EXPERIENCE

Academic Experience:

2014 - present	Assistant Professor, Director of Simulation Education, University of Central Missouri
2011 – 2013	Simulation Consultant, University of Central Missouri
2012 – 2013	Simulation Consultant, Blue Valley Center for Advanced Professional Studies
2012 – 2014	Assistant Professor, Director of Simulation, Saint Luke's College of Health Sciences
2011 - 2012	Adjunct Clinical Instructor, University of Central Missouri

Clinical Experience:

2009 – 2012	Saint Luke's East-Lee's Summit, Lee's Summit, MO Registered Nurse I, Progressive Care Unit
1996 – 1998	Clinicare Home Health Services, Kansas City, KS
1992 – 1996	University of Kansas Medical Center, Kansas City, MO, Medical Cardiac Intensive Care Unit

CAREER HIGHLIGHTS

Presentations:

- Exploring faculty preparation for debriefing using the Debriefing for Meaningful Learning Evaluation Scale (DMLES). Podium presentation. International Nursing Association for Clinical Simulation and Learning. Bradley, C.S. Grapevine, TX, Jun 15-18, 2016.
- Debriefer Evaluation with the Debriefing for Meaningful Learning Evaluation Scale (DMLES). Poster presentation. Midwest Nursing Research Society. Bradley, C.S., Dreifuerst, K.T. Milwaukee, WI, Mar 17-20, 2016. *

- The Debriefing for Meaningful Learning Evaluation Scale: Evaluating the Debriefer. Podium presentation. Professional Nurse Educator Group. Bradley, C.S. Indianapolis, IN. Oct 23-25, 2015. *

- Evaluating the Debriefer with the Debriefing for Meaningful Learning Evaluative Rubric. Podium presentation National League for Nursing Education Summit 2015. Bradley, C.S. Las Vegas, NV, Sept 30–Oct 3, 2015. *

- Pilot Testing of the Debriefing for Meaningful Learning Evaluative Rubric. Poster presentation National League for Nursing Education Summit 2015. Bradley, C.S., Dreifuerst, K.T. Las Vegas, NV, Sept 30-Oct 3, 2015. *

- Preparing the Learner for Simulation. Invited presentation. Johnson County Community College Simulation Healthcare Conference 2015. Bradley C.S. Overland Park, KS, Sept, 2015.

- Pilot Testing the Debriefing for Meaningful Learning Evaluation Scale (DMLES). Poster Presentation. Johnson County Community College Simulation Healthcare Conference 2015. Bradley, C.S., Dreifuerst, K.T. Overland Park, KS, Sept, 2015. *

- Preparing Family Nurse Practitioners for clinical exams through simulation. Poster presentation. Johnson County Community College Simulation Healthcare Conference 2015. Bradley, C.S., Skinner, K. Sept, 2015. *

- Evaluating the Debriefer with the Debriefing for Meaningful Learning Evaluation Scale. Podium presentation. Sigma Theta Tau International Research Congress 2015. Bradley, C.S. San Juan, Puerto Rico, July, 2015. *

- Preparing Family Nurse Practitioners for clinical exams through simulation. Podium presentation. Bradley, C.S., Skinner, K. International Nursing Association for Clinical Simulation in Nursing Annual Conference 2015, Atlanta, GA, June, 2015. *

- Teaching Reflective Thinking, National League for Nursing Headquarters, Debriefing Think Tank, Washington, D.C., October 29-30, 2014.

- Curriculum Integration and Best Practice in Simulation Education. Podium presentation, Simulation User Group across the Region, Bradley, C.S. Kansas City, MO, October, 2014.

- Flipped Classrooms and Active Learning Strategies. Workshop. Oklahoma Nursing Association Conference. Bradley, C.S., Tally, K.T. Oklahoma City, OK, Feb, 2014.

- Curriculum Integration of Simulation and Technology. Poster presentation. International Medical Simulation in Healthcare. Bradley, C.S., Tally, K.T., Hedrick, J. San Francisco, CA, January, 2014. *

- Trends in Nursing Education. Workshop. Saint Luke's College of Health Sciences. Bradley, C.S. October, 2013.

- Simulation as a Teaching Methodology. Simulation User Group across the Region. Bradley, C.S. August, 2012.

- Multiple Patient Leadership Scenarios. Podium presentation. Bradley, C.S., Hedrick, J. Central Regional HPSN Conference, Kansas City, MO, June 2012. *

- Exploring the Laboratory Islands. Workshop. Saint Luke's Health System. Bradley, C.S. Kansas City, MO, October, 2011
  * indicates peer reviewed

Research Activities:

- Co-investigator, Pilot Testing of the Debriefing for Meaningful Learning Evaluation Scale, Indiana University, February, 2015 – May, 2015.

- Site coordinator, Debriefing for Meaningful Learning in the Clinical Setting National research study, Saint Luke's College of Health Sciences, August, 2013 – May, 2014.

- Blinded evaluator, National Council State Boards of Nursing National Simulation Study, March, 2012 - May, 2012.

Honors, Scholarships, and Recognitions:

- National League for Nursing Mary Anne Rizzolo Doctoral Research Award, 2016
- Emily Holmquist Award presented by the Indiana University School of Nursing Alumni Association, 2016
- Jonas Scholar, Cohort VI 2015-2016

- Florence Nightingale Scholarship, Indiana University School of Nursing, 2015-2016

Outstanding/Creative Teaching Endeavors

- Professional Achievement Award, Saint Luke's College of Health Sciences, 2013
- Inductee, Saint Luke's Honor's Society, 2012

Nominations for Awards:

- Daisy Award, Saint Luke's East Lee's Summit, 2011

Service:

- Kansas City Regional Simulation Alliance, Steering Committee, 2015 - present
- Metropolitan Community College, Virtual Hospital, Advisory Board Council, 2015 - present
- UCM Department of Nursing Undergraduate Coordination Council, 2014 – present
- UCM Department of Nursing Graduate Coordination Council, 2014 – present
- UCM Department of Nursing Curriculum Committee, 2014 – present
- UCM CHST Technology Committee, 2014 – present
- UCM Department of Nursing Admissions and Progression Committee, 2015 – present
- Simulation Conference, Planning Committee, Johnson County Community College, 2014
- Simulation User Group Across the Region, 2012 - present
- Saint Luke's College of Health Sciences Curriculum Committee Co-chair, 2012-2013

CERTIFICATIONS

- BLS
- ELNEC
- Quality Matters
- Certified Nurse Educator
- Certified Simulation Health Educator

PROFESSIONAL ACTIVITIES

Professional Organizations

- American Nurse Association

- Missouri Nurse Association
- Sigma Theta Tau International, Alpha Chapter
- Golden Key International Honor Society
- Society for Simulation in Healthcare
- International Nursing Association for Clinical Simulation in Learning
- National League for Nursing
- Midwest Nursing Research Society

PUBLICATIONS

- Dreifuerst, K.T. & Bradley, C.S. (2015). Transformation of Teaching and Learning through Debriefing for Meaningful Learning. In G. Sherwood & S. Horton-Deutsch, (Eds.) *Reflective organizations: On the frontlines to transform education and practice.* Indianapolis, IN: Sigma Theta Tau International

- Bradley, C. S., & Dreifuerst, K. T. (2016). Pilot Testing the Debriefing for Meaningful Learning Evaluation Scale. *Clinical Simulation in Nursing*, *12*(7), 277-280.