2020-05-18

Measuring Learning and Engagement during Debriefing: A New Instrument

Shelly Jensen Reed
Brigham Young University - Provo, shelly-reed@byu.edu

Follow this and additional works at: https://scholarsarchive.byu.edu/facpub

Part of the Other Nursing Commons

Featured Article

Measuring Learning and Engagement during Debriefing: A New Instrument

Shelly J. Reed, PhD, DNP, APRN, CNE*

Brigham Young University, College of Nursing, Provo, UT 84602, USA

Abstract
Background: Debriefing is considered to be where the bulk of learning takes place. Examination of debriefing is needed to help determine if facilitation methods effectively contribute to the learning process. More instruments are needed to evaluate all key debriefing elements, and no tools exist that evaluate participant engagement in the learning process.
Method: Kolb's experiential learning theory was used to inform tool development. The focus of the new instrument is behaviors that indicate engagement with the learning process as the learner moves through Kolb's learning cycle. A qualitative study of debriefing facilitators was performed, and a behaviorally anchored instrument was then created. A subject matter expert review process was used to help establish validity. Reliability was evaluated in a research study looking at inter-rater consistency.
Results: The new instrument is used to count learning and engagement behaviors that are exhibited in a debriefing session.
Conclusion: The instrument is ready for use in debriefing research studies, helping evaluate participant engagement in the learning process by identifying learning and engagement behaviors that they may exhibit during a debriefing.

Debriefing, the reflective activity after an enacted simulation scenario, is considered to be a critical learning activity in simulation-based education (Dreifuerst, 2009; Palaganas, Fey, & Simon, 2016). As a core component of simulation-based learning, debriefing should be facilitated to help participants meet the objectives and needs of the learners (Alhaj Ali & Musallam, 2018; International Association of Clinical Simulation and Learning [INACSL], 2016a). Debriefing is also important in other areas of nursing education and should be used throughout a nursing curriculum to help develop reflective practitioners, as this may have a significant effect on learning outcomes (National League of Nursing [NLN], 2015).

Simulation learning is a costly and time-intensive educational process, and educators need to ensure that it is used to achieve desired outcomes (Zigmont, Kappus, & Sudikoff, 2011). As debriefing is vital to simulation learning, every method of debriefing should be evaluated to determine if the debriefing effectively contributes to

the learning process (Patterson & Klein, 2012). The question of how debriefing facilitators know a student is learning during debriefing has not been answered, as actual simulation debriefing practices in nursing education and their effect on nursing have not been well-documented, and most debriefing methods do not have associated instruments or means for evaluation (Waznonis, 2014). More recently, Alhaj Ali and Musallam (2018) found that more systematic tool development and analysis are required to evaluate simulation debriefing, focusing on all key debriefing elements. The findings of their review suggest that in addition to using established instruments, designing new instruments to evaluate debriefing is needed in nursing and health care fields, with a focus on improving intended learning outcomes (Alhaj Ali & Musallam, 2018). The purpose of this article, therefore, is to introduce a new debriefing instrument that looks at the learning process through evaluating learning and engagement behaviors occurring during simulation debriefing.

> **Key Points**
> - Debriefing is thought to be where the bulk learning takes place during simulation experiences.
> - This article describes the development of an instrument that identifies learning and engagement behaviors that may be exhibited during debriefing.
> - When combined with other debriefing evaluation instruments, it has the potential to help identify debriefing practices that contribute to participant learning.

Background

Theoretical Framework

The theoretical framework used as a basis for the new instrument is Kolb's Experiential Learning Theory (ELT). Kolb's ELT describes "the learning process as a recursive circle or spiral, where information is transferred from the teacher to the learner to be stored in declarative memory for later recall" (Kolb & Kolb, 2017, p. 15). Learning is cyclical, consisting of a concrete experience, reflective observation, abstract conceptualization, and active experimentation. The simulation scenario provides the concrete experience and is the basis for reflective observation through debriefing, where meaning is made of the experience. (Kolb & Kolb, 2009; Reed, 2012). Learning is described as a process and not in terms of outcomes. The adaptive process of learning is holistic, involving thinking, feeling, perceiving, and behaving. Knowledge is created through grasping and transforming an experience (Kolb & Kolb, 2009).

Debriefing

The most important learning component of simulation-related activities is the debriefing (Dreifuerst, 2009; Neill & Wotton, 2011; Shinnick, Woo, Horwich, & Steadman, 2011). The debriefing follows a simulation-based experience (SBE), and should be led by a trained facilitator using an evidence-based debriefing model (INACSL, 2016a). A review by Sawyer, Eppich, Brett-Fleegler, Grant, and Cheng (2016) also discusses self-debriefing as an option, and Ha and Lim (2018) identify peer-led written debriefing as an alternative to debriefing by a trained facilitator. In any type of debriefing, participants should be encouraged to reflect on and discuss the completed SBE, with feedback provided by both facilitator and participants regarding SBE performance. The purpose of debriefing is to move toward assimilation and accommodation for transfer of learning to future situations (INACSL, 2016a), and as described by ELT, to move learners through the learning cycle (Fanning & Gaba, 2007; Kolb & Kolb, 2009).

The concept of debriefing a simulation activity is not new. Steinwachs (1992) shared an article describing how debriefing should be facilitated after a simulation game, stating "debriefing is a time to reflect on and discover together what happened… and what it all means" (p. 187). Debriefing is described by Steinwachs (1992) as moving through three phases: description, analogy/analysis, and application. The description phase follows the simulation, as participants emerge from the simulated world. The purpose of this phase is to describe and express experiences and impressions and to listen to other participants to "be filled in on the whole picture." The analogy/analysis phase follows, where participants examine the simulation systematically to identify and explore parallels with real-world situations. Finally, during the application phase, participants consider what they understand from the simulation, decide what is relevant to them, and perhaps plan a course of action they would like to carry out because of these understandings. Facilitation should affirm persons who contribute to the debriefing conversation, avoid telling participants what they should have learned, respect silence, and allow adequate time for debriefing (Steinwachs, 1992).

Debriefing and Learning

Levett-Jones and Lapkin (2014) performed a systematic review regarding the effectiveness of simulation debriefing in health care research. They found improvement from pre- to post-tests with all debriefing types and found no differences in clinical or practical outcomes with the addition of video playback. Measures used in studies included in the review varied widely, with measures such as researcher-created surveys, skills checklists and knowledge examinations, and measures created for other areas outside of debriefing (Levett-Jones & Lapkin, 2014).

Other studies outside of those included in the Levett-Jones and Lapkin (2014) review provide insight into the learning that comes from debriefing. One qualitative research study explored the lived experiences of participants in simulation learning activities and concluded that the lack of consistency in the design, facilitation, and evaluation of simulations leads to confusion for both the learner and facilitator. Study conclusions did not include specific learning recommendations (Beard, 2013). A quantitative study compared postconference feedback and reflective debriefing by measuring heart failure knowledge with an instrument created from Assessment Technology Incorporated questions combined with researcher-generated questions. There was no difference in knowledge found between the two groups. The researcher concluded that there was a need for more research on debriefing in high-fidelity simulation to help determine how to best promote learning (Benhuri, 2014).

Measuring Learning Acquired during Debriefing

Kolb's ELT explains the process of experiential learning, but there is still a need for more research that can establish a cause-and-effect relationship between high-fidelity simulation and learning in undergraduate nursing education (Doolen, et al., 2016). To make the most of simulation learning, educators need to ensure that simulation is used effectively to achieve desired outcomes (Zigmont, Rappus, & Sudikoff, 2011). A study by Re (2011) concluded that simulation in conjunction with debriefing leads to attainment of learning outcomes and that further research is needed to see how those outcomes are obtained (Re, 2011).

There are a wide variety of measures used to evaluate the learning acquired during simulation debriefing, including measures such as academic self-efficacy, confidence in performance; self-assessed communication skills; and satisfaction (Lee, Kim, Kan, & Kim, 2019). Learning itself is multidimensional and is influenced by prior learning experiences. This poses additional challenges in simulation, especially when trying to attempt to isolate the learning acquired during debriefing from other simulation activities (Reed, 2016). Existing measures specific to debriefing do not measure learning, and most are specific to a debriefing method. Examples include evaluations such as the Debriefing Assessment for Simulation in Healthcare©, the Debriefing Assessment for Simulation in Healthcare Student Version©, and the Debriefing for Meaningful Learning Student Questionnaire© (Reed, 2016; Waznonis, 2014). A debriefing instrument that is nonspecific to a debriefing method is the Debriefing Experience Scale, which measures the participant experience during debriefing. While this instrument does not measure learning, it does have a subscale, "learning and making connections" where participants evaluate learning as part of their debriefing experience (Reed, 2012).

Evaluation of simulation participants using valid and reliable assessment tools can assist in achieving objectives and outcomes (INACSL Standards Committee, 2016b).

Evaluation instruments should undergo psychometric testing to yield reportable feedback, with results disseminated to build evidence-based practice guidelines for simulation debriefing (Waznonis, 2014). A study of debriefing practices showed that respondents use a variety of approaches to guide students toward meeting learning objectives. Evaluation of these approaches is lacking, yet necessary, to meet the IN-ACSL criterion recommending that debriefing focus on participant and scenario-specific objectives (Waznonis, 2015). Regarding learning as a debriefing objective, there is some evidence available on specific knowledge or skills gained during simulation and/or debriefing, but there are currently no instruments that exist that measure learning acquired during simulation debriefing (Dufrene & Young, 2014; Garden, LeFevre, Waddington, & Weller, 2015; Reed, 2016).

More research on debriefing is needed to see how learning outcomes are obtained and to determine how to best promote learning (Benhuri, 2014; Brown, 2011; Re, 2011). The use of valid and reliable instruments is necessary for the continued improvement of debriefing techniques and robust nursing faculty development (NLN, 2015). Identifying learning achieved through simulation activities is challenging for a nurse educator, as it is necessary to consider the process of knowledge discovery, enhancement of critical thinking skills, or other ways that learning can be achieved. As a result, multiple learning assessments or measures are often implemented, as it difficult to measure learning with one assessment. A standardized instrument is needed for assessing learning related to postsimulation reflection (Beard, 2013; Reed, 2016). In addition, an instrument that measures learning independent of a specific debriefing method and across simulation types will aid in the design of future simulation debriefing research that is both rigorous and feasible. This will help build the evidence for effective simulation debriefing practices. Without practical and theoretically grounded tools, simulation activities will continue to lack the element of objective assessment necessary to move evidence-based teaching practices forward (Reed, 2016; Waznonis, 2014).

Instrument Development

Behaviorally Anchored Rating Scales

Behaviorally anchored rating scales (BARS) were invented to address the deficiencies found in graphical rating-based performance scales, which have been criticized for being vague or ambiguous. The key feature of BARS is that they provide concrete behavioral examples that raters can use during evaluation (Klieger et al., 2018). Examples of BARS include employment interview performance, personality traits, classroom teamwork, and motivation. BARS provide many evaluative advantages, including a reduction of

construct-irrelevant variance in appraisal based on their emphasis on specific, concrete, observable behaviors as a means of defining the dimensions to be judged. Consistent with this reasoning, some studies demonstrate that ratings using BARS usually exhibit less measurement bias than those produced using other types of scales (Klieger et al., 2018). As behaving is part of the adaptive learning process described by Kolb's ELT (Kolb & Kolb, 2009), the development of a behaviorally anchored rating scale is appropriate to measure behavioral aspects of the learning process.

Evidence of Content Validity

In instrument development, the first step in the design of any measure it is to clarify the purpose for the measure (Waltz, Strickland, & Lenz, 2010). The purpose for the planned instrument is to provide an objective measure of the learning process during debriefing; however, it was unknown how debriefing facilitators know learning is happening. Approaches to qualitative research seek to arrive at an understanding of a particular phenomenon from the perspective of those experiencing it (Vaismoradi, Turunen, & Bondas, 2013), with the end goal as "a rich contextualized understanding of some aspect of the human experience through an intensive study of particular cases" (Polit & Beck, 2010, p. 1451). Thus, a qualitative descriptive research study was conducted to describe how debriefing facilitators identify participants are learning during simulation debriefing (Reed, 2016). The study also helped meet procedures for BARS development (Klieger, et al., 2008), where subject-matter experts (SMEs) (in this case, debriefing facilitators) provide examples of behaviors.

Content validity in instrument development is largely a matter of judgement, consisting of two distinct phases. The first is "a priori efforts by the scale developer to enhance content validity though careful conceptualization and domain analysis prior to item generation" (Polit & Beck, 2006, p. 490). The qualitative research study described meets criteria for this phase. The second phase is posteriori efforts evaluating scale content through expert assessment (Polit & Beck, 2006). This was accomplished through a content validity index (CVI) calculation for both items, subscales, and the overall scale. Three to ten experts are recommended for the calculation, and six experts completed the review of this instrument. The expert reviewers rated items as "not relevant, somewhat relevant, quite relevant, or highly relevant," with an item content validity index (I-CVI) computed from the number of experts rating it as "quite relevant" or "highly relevant" (Polit & Beck, 2006). Statistical results were calculated using IBM SPSS Statistics, version 25. The scale content validity index (S-CVI) computed at 0.87 using the average of I-CVI. "Many [methods] writers have indicated that an S-CVI of 0.80 or higher is acceptable" (Polit & Beck, 2006, p. 491), with the 0.87 falling above acceptability criteria. While

researchers do not usually provide I-CVI values (Polit & Beck, 2006), the item and subscale (construct) results are provided in Appendix A for transparency purposes. This is also to avoid only reporting S-CVI, which according to Polit & Beck, 2006, p. 491, "is where the problems lie." Twenty-one of 26 items have an I-CVI of 0.83 or greater, 0.83 is the recommended acceptability rating for I-CVI for a six-expert review (Polit, Beck, & Owen, 2007). Although the five items were below the 0.83 rating, they were not eliminated from the instrument. This was due to the acceptability of the overall S-CVI, and inclusion of these items as a result of both robust construct validity and a-priori content validity efforts.

Evidence of Construct Validity

"Constructs are intangible collections of abstract concepts and principles which are inferred from behavior and explained by educational or psychological theory" (Downing, 2003, p. 831). Items for new instruments can be developed through literature search, seeking expert opinions, and population sampling as the researcher defines the construct of interest and its dimensions (DeVon, et al., 2007). Constructs for the new instrument were developed from the behavioral themes identified from a qualitative study of debriefing facilitators, andprovided the learning and engagement behaviors (items) for the new instrument. This meets another procedure during BARS development, where performance categories are created and labeled by content similarities (Klieger, et al., 2018).

Three templates for a new instrument were developed from the constructs; these were reviewed by over 50 simulation experts at the Great Researcher session at the 2017 INACSL conference. The expert researchers gave feedback on the three versions, withall suggestions and feedback from these experts incorporated into the one version of instrument that was found most favorable to the researchers. This instrument was used in a pilot study comparing debriefing sessions. Study results were used to further refine the instrument, incorporating feedback given by the research team who used it in the study. Refinements included collapsing duplicate items, and reformatting the instrument for easier use. An example of reformatting included putting all items on a single page, rather than multiple pages. This two-step SME review process, following the initial SME provided by the qualitative study, helps meet the retranslation component of BARS development. Retranslation is where a second group of SMEs places behaviors into the performance category in which they believe it best fits, and is followed by a review of yet another different group of SMEs (Klieger, et al., 2018).

Evidence of Response Process Validity

Response process validity questions whether those completing an instrument understand the constructs of a

survey as intended by the researchers. This validity cannot be tested statistically and is examined through respondent observation, interviews, and feedback, accomplished for this instrument through the previously described three-level SME review process (Yuhas & BrckaLorenz, 2018).

Evidence for Validity and Consequences of Testing

This evidence looks at the proposed interpretation of data collected by the instrument, and intended and unintended consequences of testing (American Educational Research Association [AREA], American Psychological Association [APA], and National Council on Measurement in Education [NCME], 2014). Based on learning as accomplished through the learning cycle described by Kolb's ELT, the instrument is intended to measure debriefing learning and engagement behaviors in undergraduate nursing students. In addition to using it in a different population without additional testing, unintended consequences could result when using it to evaluate other aspects of debriefing, such as the student experience, or as an evaluation of the debriefer, where other established instruments provide more appropriate measures.

Reliability

The instrument includes a list of statements or behaviors that a debriefing participant can make. Data collected by the instrument are count data, meaning the statements or behaviors are coded into only one category and are thus mutually exclusive. Each statement or behavior is tallied and does not have a nominal (or ordinal) assessment to it. For example, if a debriefing participant "comments on peers/group performance" (sharing with peers category), a tally is given but is not assigned a scaling such as "low," "medium," or "high." Measures of reliability such as a Cronbach's alpha cannot be calculated from count data, and thus this value is not provided for this instrument.

Broadened measurements over the past few decades include areas that are otherwise difficult to assess. The flexibility in assessment has increased variations in testing replication, and subsequently increased measurement error, a component of reliability (AREA, APA, & NCME, 2014). "However, some of these sacrifices in reliability/ precision may reduce construct irrelevance or construct underrepresentation, and thereby improve the validity of these scores … [and] can sometimes provide more direct measures of the attribute of interest" (AREA, APA, & NCME, 2014, p. 36).

Reliability data on test scoring using subjective judgement should provide inter-rater consistency in scoring, clearly describing data based on raters' performance (AREA, APA, & NCME, 2014). For this instrument, a research study comparing debriefing sessions of undergraduate nursing student groups was used to look at inter-rater agreement. Before data collection during the debriefing sessions, raters from the research team watched a video-recorded debriefing together and discussed with each other their reasons for instrument scoring. The video-recording was paused following a statement or behavior, with research team members continuing on after agreement on scoring was reached. Team members then used the instrument to evaluate the group debriefs of eleven different simulations. IBM SPSS Statistics, version 25 was used to calculate the results. Inter-rater agreement of the two evaluators was within 4.5% in the eleven paired debriefing observations. Four of the eleven observations had 100% inter-rater agreement.

Instrument Description

The new instrument is behaviorally anchored and is titled "Debriefing Engagement and Learning Behaviors (DELB)." Instrument anchors (constructs) include critical review of actions, connecting learning, planning for future experiences, excitement and engagement, and direct expressions of learning. Behaviors are identified for each anchor. For example, a behavior for the connecting learning anchor is "gives solutions connecting past and present learning." The instrument is contained on one page and includes demographic information at the top. Beneath the demographic information are three columns. The first column contains the constructs (behavioral anchors), the second column contains the behaviors, and the third column is empty for observer scoring. Items and subscales are included in Appendix A.

Scoring

Two or more raters use the instrument to score a debriefing. To establish consistency before official use, these raters should watch a video-recorded debriefing session together, discussing with each other what they are seeing, and then assigning a tally (mark) for an observed behavior to a behavior on the instrument. This observation/discussion should continue until consensus is reached on instrument scoring for observed behaviors. The raters are then ready to use the instrument when observing an actual debriefing session. The debriefing session is timed from the beginning to end of the session, to determine the overall debriefing time in minutes. Raters observe the debriefing and tally separately behaviors observed in the session. Each behavior receives one point, based on the participant behavior it most closely matches. On completion of the debriefing, total points are calculated for all raters and then divided by the number of raters. These points are then divided by the total minutes for the session, resulting in a behavior/minute score. Inter-rater agreement is calculated comparing the scores of the raters.

Behavioral counts are obtained from the completed instrument. Count data are the number of occurrences of

a behavior in a fixed period. The coding of a behavior to just one area on the instrument leaves the possibility of some listed behaviors having low or no counts. Areas with a low arithmetic mean require analysis using regression models such as Poisson regression to provide an appropriate analysis for count data (Coxe, West, & Aiken, 2009).

Discussion

The relationship created between the teacher, the learner, and the subject matter is the crux of experiential learning. ELT places the subject to be learned in the center, and both teacher and learner experience it, leveling their learner/teacher relationship. The reflective observation phase of the learning cycle provides a venue for sharing learning by both learners and teacher, as all have a perspective on the subject. In fact, all modes of the learning cycle are experiences, including the reflective observation provided by debriefing (Kolb & Kolb, 2017). ''The learning cycle is driven by the integration of action and reflection and experience and concept'' (Kolb & Kolb, 2017, p. 14). While the new instrument does not achieve the original objective of measuring learning, it does focus on debriefing participant learning and engagement behaviors, evidence of involvement in the learning process as described by Kolb's learning cycle.

Regarding instrument content validity, five of 26 items had I-CVI scores below the recommended 0.83. Constructs and items for the instrument were generated from themes of a qualitative study of facilitators who debrief undergraduate nursing students, providing a solid foundation for their inclusion in the instrument. One expert reviewer evaluating content validity is not from the country where the qualitative study was conducted. A second expert reviewer facilitates debriefs in a hospital environment with primarily team-based simulations, which is a different practice setting than the debriefing facilitators included in the qualitative study. These differences may have affected I-CVI ratings. This raises the issue of generalizability in using the instrument to evaluate debriefing outside of those for undergraduate nursing students; it should not be used in other populations without first performing reliability and validity testing in the population of interest.

The development of the instrument using a three-step review by SMEs provides robust construct validity. As a behaviorally anchored scale, the instrument has the potential to provide a more objective measure of engagement in the learning process, as compared with other measures of learning used in prior debriefing studies. Learning is multifaceted and includes prior experiences, attitudes, knowledge, and even test taking or instrument completion variants, all influences coming from outside of a debrief which may impact the score on a participant-completed measure. Another strength of the instrument is that it is not attached to any established debriefing method and is thus not biased for or against any one method.

Specific uses where the instrument may be helpful include some areas recommended for future research by Doolen et al. (2016), including the effect of debriefing strategies on learning outcomes of simulation, and inclusion of experimental studies with a more rigorous design. Another area for possible use is comparing debriefing methods and the context in which use of one debriefing method may be more appropriate than another (Dufrene & Young, 2014; Sawyer et al., 2016). The instrument could also be used to help examine debriefing components, for example, length of debriefing and other time-related issues, environment, use of debriefing adjuncts such as video during debriefing, discussion of simulation observation sheets in debriefing, and co-facilitation (Reed, 2016; Waznonis, 2016).

Limitations

This instrument only measures learning and engagement behaviors and cannot measure what a participant learns during debriefing. As both teacher (facilitator) and student interact with the subject matter in ELT (Kolb & Kolb, 2017), adding additional instruments that measure their learning and engagement can help provide a more comprehensive look at debriefing. Another limitation is that reliability and validity for the scale was established when debriefing undergraduate nursing students, and therefore, reliability and validity would need to be established when using the instrument with other populations.

Conclusion

The new instrument measures learning and engagement behavior exhibited as debriefing participants move through the learning cycle. When combined with other existing debriefing measures, it has the potential to contribute to a simulation debriefing assessment that is both rigorous and feasible. It has the potential to help build the evidence determining effective simulation debriefing practices, (Waznonis, 2014) including the practices that contribute to participant learning. Using appropriate evaluation can lead to improved simulation experiences, including meeting learning outcomes (INACSL Standards Committee, 2016a). With the possibility of increasing learning, continued investigation of simulation debriefing seems well worth the effort.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecns.2020.03.002.

References

Alhaj Ali, A., & Musallam, E. (2018). Debriefing quality evaluation in nursing simulation-based education: An integrative review. *Clinical Simulation in Nursing*, *16*, 15-24. https://doi.org/10.1016/j.ecns.2017.09.009.

American Educational Research AssociationAmerican Psychological AssociationNational Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Beard, R. (2013). Exploring the lived experiences of participants in simulation-based learning activities. (Doctoral Dissertation, Linwood University). In *ProQuest Dissertations and Theses database. (UMI No. 3605518)*.

Benhuri, G. (2014). Debriefing after human patient simulation and nursing students' learning. (Doctoral Dissertation, University of Phoenix). In *ProQuest Dissertations and Theses database. (UMI No. 3585955)*.

Brown, F. S. (2011). Nursing faculty beliefs and practices regarding debriefing human patient simulation experiences. (Doctoral Dissertation, University of California, Davis). In *ProQuest Dissertations and Theses database. (UMI No. 3474352)*.

Coxe, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: A gentle introduction to poisson regression and its alternatives. *Journal of Personality Assessment*, *91*(2), 121-136. https://doi.org/10.1080/00223890802634175.

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

Doolen, J., Mariani, B., Atz, T., Horsley, T. L., O'Rourke, J., McAfee, K., & Cross, C. L. (2016). High-fidelity simulation in undergraduate nursing education: A review of simulation reviews. *Clinical Simulation in Nursing*, *12*(7), 290-302. https://doi.org/10.1016/j.ecns.2016.01.009.

Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education*, *37*, 830-837.

Dreifuerst, K. T. (2009). The essentials of debriefing in simulation learning: A concept analysis. *Nursing Education Perspectives*, *20*(2), 109-114.

Dufrene, C., & Young, A. (2014). Successful debriefing—best methods to achieve positive learning outcomes: A literature review. *Nurse Education Today*, *34*, 372-376.

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

Garden, A. L., LeFevre, H. L., Waddington, J. M., & Weller, J. M. (2015). Debriefing after simulation-based non-technical skill training in healthcare: A systemic review of effective practice. *Anaesthesia & Intensive Care*, *43*(3), 300-308.

Ha, E. H., & Lim, E. J. (2018). Peer-led debriefing versus instructor-led debriefing: Using multimode simulation. *Clinical Simulation in Nursing*, *18*, 38-46.

INACSL Standards Committee. (2016a). INACSL standards of best practice: SimulationSM simulation glossary. *Clinical Simulation in Nursing*, *12*(S), S39-S47. https://doi.org/10.1016/j.ecns.2016.09.012.

INACSL Standards Committee. (2016b). INACSL standards of best practice: Simulation participant evaluation. *Clinical Simulation in Nursing*, *12*(S), S26-S29.

Klieger, D. M., Kell, H. J., Rikoon, S., Burkander, K. N., Bochenek, J. L., & Shore, J. R. (2018). *Development of the behaviorally anchored rating scales for the skills demonstration and progression guide. (Educational Testing Service Research Report No. RR-18-24)*. Princeton, NJ: Educational Testing Service.

Kolb, A. Y., & Kolb, D. (2009). Experiential learning theory: A dynamic, holistic approach to management learning, education, and development. In Armstrong, S. J., & Fukami, C. V. (Eds.), *The SAGE handbook of management learning, education, & development*. Los Angeles, CA: Sage Publications. (pp. 42-68).

Kolb, A. Y., & Kolb, D. A. (2017). Experiential Learning Theory as a guide for experiential educators in higher education. *ELTHE: A Journal for Engaged Educators*, *1*(1), 7-44.

Lee, M.-N., Kim, S.-J., Kang, K.-A., & Kim, S. (2019). Comparing the learning effects of debriefing modalities for the care of premature infants. *Nursing and Health Sciences* 1-11. https://doi.org/10.1111/nhs.12662.

Levett-Jones, T., & Lapkin, S. (2014). A systematic review of the effectiveness of simulation debriefing in health professional education. *Nurse Education Today*, *34*, e58-e63.

National League of Nursing [NLN]. (2015). *NLN Vision Series: Debriefing across the curriculum*. Retrieved from http://www.nln.org/docs/default-source/about/nln-vision-series-%28position-statements%29/nln-vision-debriefing-across-the-curriculum.pdf?sfvrsn=0.

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

Palaganas, J. C., Fey, M., & Simon, R. (2016). Structured debriefing in simulation-based education. *AACN Advanced Critical Care*, *27*(1), 78-85.

Patterson, B. J., & Klein, J. M. (2012). Evidence for teaching: What are faculty using? *Nursing Education Perspectives*, *33*(4), 240-245.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critiques and recommendations. *Research in Nursing & Health*, *29*, 489-497.

Polit, D. F., & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. *International Journal of Nursing Studies*, *47*, 1451-1458. https://doi.org/10.1016/j.ijnurstu.2010.06.004.

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, *30*, 459-467.

Re, A. (2011). The impact of human patient simulation on the attainment of learning outcomes. (Doctoral Dissertation). In *ProQuest Dissertations and Theses database. (UMI No. 3472945)*.

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. https://doi.org/10.1016/j.ecns.2011.11.002.

Reed, S. J. (2016). Identifying learning acquired during debriefing. (Doctoral dissertation). In *UNLV Theses, Dissertations, Professional Papers, and Capstones*. Retrieved from http://digitalscholarship.unlv.edu/thesesdissertations/2724.

Sawyer, T., Eppich, W., Brett-Fleegler, M., Grant, V., & Cheng, A. (2016). More than one way to debrief: A critical review of healthcare simulation debriefing methods. *Simulation in Healthcare*, *11*, 209-217.

Shinnick, M. A., Woo, M., Horwich, T. B., & Steadman, R. (2011). Debriefing: The most important component in simulation? *Clinical Simulation in Nursing*, *7*, e105-e111. https://doi.org/10.1016/j.ecns.2010.11.005.

Steinwachs, B. (1992). How to facilitate a debriefing. *Simulation and Gaming*, *23*(2), 186-195.

Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing and Health Sciences*, *15*, 398-405.

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

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), 110-119.

Waznonis, A. (2016). Faculty descriptions of simulation debriefing in traditional baccalaureate nursing programs. *Nursing Education Perspectives*, *37*(5), 262-268. https://doi.org/10.1097/01.NEP.0000000000000065.

Yuhas, B., & BrckaLorenz, A. (2018). *Response process validity: 2013-2016 write-in response analysis. FSSE Psychometric Portfolio*. Retrieved from fsse.indiana.edu.

Zigmont, J. J., Kappus, L. J., & Sudkioff, S. N. (2011). Theoretical foundations of learning through simulation. *Seminars in Perinatology, 35*, 2.