Clinical Simulation in Nursing (2013) 9, e393-e400

Clinical Simulation in Nursing

www.elsevier.com/locate/ecsn

Featured Article

An Updated Review of Published Simulation Evaluation Instruments

Katie Anne Adamson, PhD, RN[a,*], Suzan Kardong-Edgren, PhD, RN[b], Janet Willhaus, MSN, RN[c]

[a]*University of Washington Tacoma, Tacoma, WA 98402-3100, USA*
[b]*Boise State University, WA, USA*
[c]*Washington State University, Spokane, WA 99210, USA*

KEYWORDS
evaluation;
instrument;
simulation

Abstract: Interest in simulation as a teaching and evaluation strategy in nursing education continues to grow. Mirroring this growth, we have seen a proliferation of instruments designed to evaluate simulation participant performance. This article describes two frameworks for categorizing simulation evaluation strategies and provides a review of recent simulation evaluation instruments. The review focuses on four instruments that have been used extensively in the literature, objective structured clinical examinations (OSCE's) including four OSCE instruments, and an extensive list of new instruments for simulation evaluation.

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

Simulation use continues to grow and develop in nursing and other programs educating health care providers around the world. DeVita (2009) argues that simulation should be a core educational strategy because it is "measurable, focused, reproducible, mass producible, and importantly, very memorable" (p. 46). Both the National Council of State Boards of Nursing and the National League for Nursing are conducting research about the use of simulation as a teaching and evaluation method (Hayden, 2011; Rizzolo, Oermann, Jeffries, & Kardong-Edgren, 2011). However, Tanner (2011) recently noted how "little investment there has been in developing suitable measures for the assessment of learning outcomes, particularly those relevant for a practice discipline" (p. 491). Recent reviews of the literature in nursing (Davis & Kimble, 2011; Yuan, Williams, Fang, & Ye, 2012), pharmacy (Bray, Schwartz, Odegard, Hammer, & Seybert, 2011) and medicine (Kogan, Holmboe, & Hauer, 2009) echo a continued quest for meaningful ways to evaluate participants in simulation activities.

In response to repeated requests for an updated and expanded list of evaluation instruments, this article provides a follow-up to the original instrument review article (Kardong-Edgren, Adamson, & Fitzgerald, 2010). The purposes for this article include (a) discussing existing frameworks for categorizing simulation evaluation strategies and (b) using an adaptation of these frameworks to provide the following:

1. An update on four instruments from our original review that have been cited extensively in the literature

* Corresponding author: kadamson@u.washington.edu (K. A. Adamson).

1876-1399/$ - see front matter © 2013 International Nursing Association for Clinical Simulation and Learning. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ecns.2012.09.004

2. A review of objective structured clinical examinations (OSCEs), including the development of four OSCE instruments in undergraduate nursing education
3. A report on instruments that are either new or were not included in the original instrument review article (Kardong-Edgren, et al., 2010) and that are appropriate for simulation evaluation

Key Points
- Translational Science and Kirkpatrick's (1994) Levels of evaluation are two useful frameworks for categorizing simulation evaluations.
- Existing evaluation instruments continue to be adapted and applied.
- New evaluation instruments continue to emerge.

Frameworks for Categorizing Evaluation Strategies

Two useful frameworks that have emerged to categorize various evaluation strategies are translational science research (TSR; McGaghie, Draycott, Dunn, Lopez, & Stefanidis, 2011) and Kirkpatrick's (1994) levels of evaluation. The following is a brief overview of these frameworks, which will be used to categorize instruments in following sections.

TSR

The National Institutes of Health (2011, 2012) describe translational research as a continuum on which scientific discoveries move from preclinical (or bench) research to practical applications in patient care at the bedside and ultimately affect health care outcomes. In short, TSR can be thought of as research that takes new knowledge from "bench to bedside and beyond." Nomenclature in the field of TSR is somewhat contested (Woolf, 2008). However, the concept is highly applicable to simulation evaluation research. For the purposes of this article, we are adopting the overview provided by Dougherty and Conway (2008) and applied to simulation evaluation by McGaghie et al. (2011).

Translation Phase 1 designates preclinical activities (Woolf, 2008) that are meant to assess the efficacy of care. Relating this to simulation, we might say that this level of research demonstrates, in the simulation lab, whether students have learned something. Translation Phase 2 designates activities meant to assess who benefits from care. Relating this to simulation, we might say that these activities demonstrate whether what students learned in the simulation lab carries over to the actual patient care setting. Finally, Translation Phase 3 designates activities that are meant to assess whether improved care yields improved outcomes in the broader health care arena. Relating this to simulation, we might say that these activities demonstrate whether what was learned in the simulation lab and demonstrated in the patient care setting results in improved health outcomes. Phases 1 to 3 help describe the quality and applicability of simulation evaluation activities, with Phase 3 activities being the pinnacle of research because they describe how simulation affects health outcomes.

Kirkpatrick's Levels of Evaluation

In a similar fashion, Kirkpatrick's (1994) four levels of evaluation are helpful in describing what type of evidence different simulation evaluation strategies produce. The four levels, reaction, learning, behavior, and outcomes, are described in Figure 1, using language from Boulet Jeffries, Hatala, Korndorffer, Feinstein, & Roche, (2011, p. S50), along with the corresponding TSR nomenclature. In this combination of the TSR and Kirkpatrick frameworks for describing types of simulation evaluation evidence, learning at Level 2 (Translation Phase 1) may be subdivided into affective, cognitive, and psychomotor learning. Also, Kirkpatrick's Level 1, reaction, is not applicable to translational research.

Review of Simulation Evaluation Instruments

The following sections include an update regarding four instruments from our original review, a review of OSCEs, and a report on instruments that either are new or were not included in the original instrument review article. They refer to Tables 1 through 3, which indicate TSR phases and Kirkpatrick's levels.

Update on Instruments From Original Review

Four instruments from our original review have been used repeatedly to evaluate performance and learning and are reported in the literature: the Sweeny-Clark Simulation Performance Evaluation Tool (Clark (2006) Tool©); the Clinical Simulation Evaluation Tool (CSET; Radhakrishnan, Roche, & Cunningham, 2007); the Lasater Clinical Judgment Rubric (LCJR©; Lasater, 2007); and the Creighton Simulation Evaluation Instrument (C-SEI™; Todd, Manz, Hawkins, Parsons, & Hercinger, 2008), subsequently modified to create the Creighton Competency Evaluation Instrument (Hayden, Kardong-Edgren, Smiley, & Keegan, Reliability and validity testing of the Creighton Competency Evaluation Instrument [CCEI] for use in the NCSBN National Simulation Study [2012 unpublished data]). Table 1 provides an overview of each of these instruments.

The Sweeny-Clark Simulation Performance Evaluation Tool (Clark (2006) Tool©) has been modified for specific evaluation needs, but the authors have not described additional reliability or validity assessments of these modifications. Similarly, the CSET, developed and originally published by Radhakrishnan et al. (2007), has been modified to suit diverse performance evaluation needs, including simulated and standardized patient encounters involving congestive heart failure (Adamson, in press), myocardial

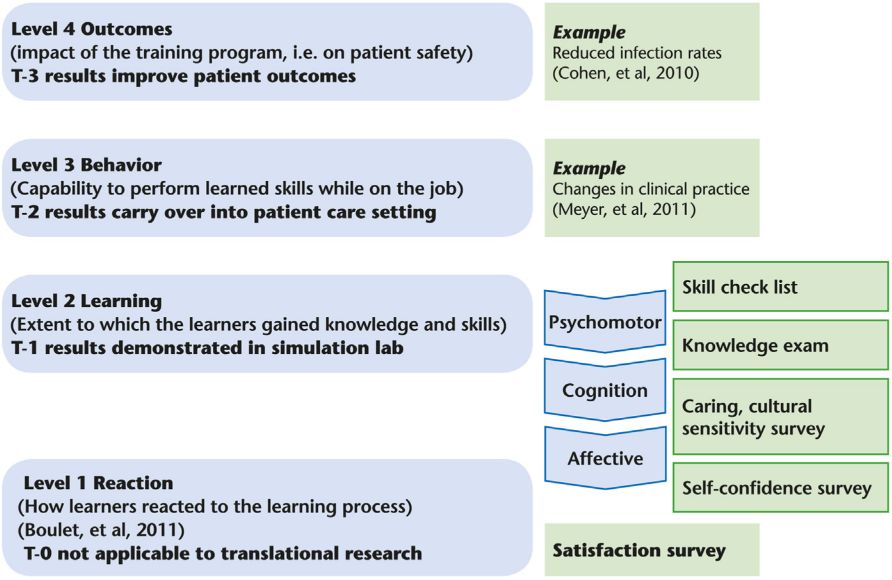

Figure 1 T-1, Translation Phase 1; T-2, Translation Phase 2; T-3, Translation Phase 3; T-0, not applicable to translational research.

infarction, and a chest wound (Grant, Moss, Epps, & Watts, 2010). Grant et al. (2010) report interrater reliability findings from their modification of the CSET. The LCJR (Lasater, 2007) has been used for a variety of purposes, including debriefing (Mariani, Cantrell, Meakim, Prieto, & Dreifuerst, in press) and evaluation of technical skills such as IV insertion (Reinhardt, Mullins, De Blieck, & Schultz, 2012). Furthermore, Adamson, Gubrud, Sideras, and Lasater (2012) reported extensive reliability and validity findings from a range of studies used to assess the psychometric properties of the LCJR. Finally, the C-SEI™, originally developed and published by Todd et al. (2008), has undergone extensive reliability and validity assessments and was adapted by Hayden et al. (April 2012, unpublished data) for use in the National Council of State Boards of Nursing Simulation Study.

Each of the instruments in Table 1, as well as the authors who chose to use existing tools rather than develop their own, represents a concerted effort to build on previous knowledge. In our original instrument review article, we underscored the importance of ‘‘further use and development of these published simulation evaluation instruments’’ (Kardong-Edgren et al., 2010, p. e33). The continued use and psychometric assessments of the Sweeny-Clark Simulation Performance Evaluation Tool, CSET, LCJR, and C-SEI are exemplars of this effort.

Instruments for OSCEs

An OSCE is a set of performance-based scenarios in which students may be observed in the demonstration of clinical behaviors (McWilliam & Botwinski, 2010). A series of brief simulation work stations are developed for evaluation of specific tasks and behaviors. Students rotate through each station, where performance is assessed with checklists and rating scales. Stations may include the use of task trainers, standardized patients, manikins, screen-based simulation, and role playing. Four of the articles reviewed for this update focused on the use of the OSCE to evaluate performance. Table 2 provides an overview of each of these instruments.

The development of an OSCE for psychiatric nursing (Selim, Ramadan, El-Gueneidy, & Gaafer, 2012) details the time and effort required to ensure that each of 11 stations demonstrated content validity and reliability. Interrater reliability of scoring tools at each station is reported. This OSCE was developed for nursing students in Egypt, but a replication of this method could further development of valid and reliable instruments in nursing programs worldwide. There are four essential areas for consideration in the development of OSCE reliability and validity: measuring context-reliant competence, measuring competence versus performance, measuring professional behavior, and measuring integration of skills (Mitchell, Henderson, Groves, Dalton, & Nulty, 2009).

Multiple short-duration stations where a technical element can be observed are recommended when OSCEs are used. Hutton et al. (2010) compared a computer-based simulation examination with a parallel OSCE examination. Although students were able to assess conceptual and calculation skills with the computer evaluation instrument, technical skill errors were identified primarily by the OSCE. Another report using an OSCE for evaluation of medication skills in pediatrics (Cazzell & Howe, 2012) emphasizes the importance of interrater reliability in evaluating performance.

Table 1 Updates on Instruments from Original Review Article

Articles: Original; Subsequent Publications Related to the Instrument	Instrument	Reliability and Validity	Kirkpatrick and TSR	Special Notes
Original article, Clark, 2006, p. e76	The Sweeny-Clark Simulation Performance Evaluation Tool (Clark Tool©): author designed and copyrighted		K2 T1	
Gantt, 2010	Included rubric in article	"The rubric was found to be a practical tool that could potentially be used with or without skills checklists." (p. 101)	K2 T1	Tool used to evaluate undergraduate nursing students in manikin-based simulation; discussion included considerations for using tool for group evaluations.
Adamson, in press	Modified rubric	No additional reliability or validity data provided	K2 T1	Author modified tool to evaluate undergraduate nursing students in standardized patient encounters and combined scores with additional performance-evaluation measures.
Original article, Radhakrishnan, Roche, & Cunningham, 2007	CSET included in article	Reliability and validity not reported	K2 T1	Instrument included in article
Grant, Moss, Epps, & Watts, 2010 Grant, Epps, Moss, & Watts, 2009	Observational data collection tool, adapted from the CSET	"Fleiss's kappa coefficients used to assess interrater reliability among data collectors ranged from .71 to .94. Percentage agreement among data collectors ranged from 85% to 97%."	K2 T1	Tool was modified for use evaluating participants "caring for two patients; one with a myocardial infarction and one with a stab wound to the chest." (p. e181) Contact information for primary author of Grant, Moss, Epps & Watts (2010): grantj@uab.edu
Original article, Lasater, 2007	LCJR	Original report describes reliability and validity assessments under way.	K2 T1	Instrument based on Tanner's Model of Clinical Judgment
Mariani, Cantrell, Meakim, Prieto, & Dreifuerst, in press	Used LCJR with Debriefing for Meaningful Learning method	In this report α ranged from .80 to .97. Reports reliability and internal consistency from other studies within the article	K2 T1	LCJR was used to compare structured and unstructured debriefing methods. No significant difference was found in scores between structured and unstructured debriefing.
Adamson, Gubrud, Sideras, & Lasater, 2011)	Used original rubric	Three methods for validity and reliability assessment are described, including results.	K2 T1	
Reinhardt, Mullins, De Blieck, & Schultz, 2012	Adapted LCJR for IV skill	Did not report reliability and validity of this adaptation	K2 T1	

Original article, Todd et al., 2008	C-SEI© (tool being used in multisite national studies. See text)	Original pilot reports interrater reliability of 84.4% to 89.9%.	K2 T1	Materials available to orient faculty in an effort to improve consistency in evaluation
Manz, Hercinger, Todd, Hawkins, & Parsons, in press		Additional data about applications of the C-SEI		
Adamson, Parsons, Hawkins, Manz, Todd, & Hercinger, 2011		Further information about the reliability of data from the C-SEI		
Hayden, Kardong-Edgren, Smiley, & Keegan, (April 2012, unpublished data)	C-CEI	Extensive reliability and validity testing completed and findings reported	K3 T2	Update of the CSEI with QSEN language; tested in both simulation and clinical; 23-item objective evaluation instrument

C-CEI = Creighton Competency Evaluation Instrument; C-SEI© = Creighton Simulation Evaluation Instrument; K2 = Kirkpatrick Level 2; K3 = Kirkpatrick Level 3; LCJR = Lasater Clinical Judgment Rubric©; QSEN = Quality and Safety Education for Nurses; T1 = Translation Phase 1; T2 = Translation Phase 2; TSR = translational science research.

New and Previously Unreported Instruments

Since the publication of our original review article, there has been a sharp increase in new simulation evaluation instruments in the literature (Kardong-Edgren et al., 2010). A sampling of these instruments and citations for the articles that cited them are included as Table 3 (view online extra at www.nursingsimulation.org). Several trends and other noteworthy information in the table deserve mention here.

Two articles cited in the table used the Spielberger State-Trait Anxiety Inventory to evaluate participant anxiety related to simulation activities (Gantt, in press; Gore, Hunt, Parker, & Raines, 2011). This represents an interesting exploration of the reactions of participants and the authenticity of their emotional responses related to simulated patient encounters. Additional research is under way about the biological markers related to stress and anxiety experienced by participants in simulation.

The National League for Nursing's Simulation Design and Student Satisfaction and Self-Confidence in Learning scales (Jeffries & Rizzolo, 2006) continue to be popular (Adamson, in press; Prentice, Taplay, Horsley, Payeur-Grenier, & Delford, 2011; Swenty & Eggleston, 2011). These, like most simulation evaluation instruments, focus on low-level learner reaction and learning (Kirkpatrick's Levels 1 and 2 and TSR Phase 1). Within the category of learning, most evaluation instruments focus on cognitive learning. This is disappointing because these low levels of evaluation may not reflect the effects simulation training has on the most important stakeholders in health care education: the patients.

It is not necessarily suprising that most of the instruments included in this review focus on reaction and learning. As Hemman, Gillingham, Allison, and Adams (2007) describe, it is very difficult to validate performance-based evaluation instruments. A determination of competency is often subject to the experience, perception, training, and knowledge of the evaluator. Hemman et al. (2007) also point out that simulation testing relies on the ability of an evaluator to create, manipulate, and control conditions so that the person being tested can demonstrate skills and performance. Reaction and learning are often the low-hanging fruit of simulation evaluation. However, we would like to challenge simulation practitioners and scholars to aspire to higher levels of evaluation that reflect how simulation training affects participants' behaviors and patient outcomes (Kirkpatrick's Levels 3 and 4 and TSR Phases 2 and 3).

Discussion

We believe that this menu of simulation evaluation instruments can be very useful for educators looking for ways to evaluate simulation participants. However, like any resource, it is subject to misuse. One caution we would like to extend is about the specificity of reliability and validity data. When

Table 2 Objective Structured Clinical Examinations

Articles, Original and Subsequent	Instrument	Reliability and Validity	Kirkpatrick and TSR	Special Notes
Cazzell & Howe, 2012	OSCE	Reports the interrater reliability for this single simulation with numerous technical skill and nursing judgment components	K2 T1	Use of a single simulation to evaluate various skills of first-semester senior nursing students in a pediatric simulation
Mitchell, Henderson, Groves, Dalton, & Nulty, 2009	OSCE	High face validity, but can result in inconsistencies.	K2 T1	Recommends to have larger numbers of short stations lasting no more than 5 minutes; discusses four essential areas to consider in evaluation.
Hutton et al., 2010	OSCE for medication calculations and a computerized assessment	Internal consistency reliability of the computer assessment reported. OSCE tasks were based on the computer test, but with psychomotor manipulation of tablets, syringes, IV administration materials, and liquid medicines.	K2 T1	Compared a computer-based assessment with an OSCE for medication administration. Although computer model was able to assess conceptual and calculation skills, it was not able to accurately assess psychomotor skills or degrees of accuracy with actual medication preparation.
Selim, Ramadan, El-Gueneidy, & Gaafer, 2012	OSCE	Reports process for developing validity and reliability for an OSCE in psychiatric nursing.	K2 T1	Authors spent 4 months developing a valid and reliable OSCE format for nursing students in Alexandria, Egypt, which had 11 stations with two evaluators at each station.

K2 = Kirkpatrick Level 2; OSCE = Objective Structured Clinical Examination; T1 = Translation Phase 1; TSR = translational science research.

educators are selecting an instrument for use in performance evaluation, it is not enough to select a tool from a list with high marks reported in reliability and validity. It is important to consider whether the instrument is appropriate for the population and the activity to which it is being applied.

In research, when an instrument is used in a new population or for a measurement purpose different from what was originally intended, the researchers should report the process and statistics associated with validating the instrument for the new purpose. Using an instrument to evaluate populations and purposes beyond the original intent is like trying to measure a cup of milk with a yardstick. Although it is possible, without accurate knowledge about the vessel for the liquid, it would be difficult to determine whether the amount of milk really equaled 1 cup. Before an instrument is used to evaluate student performance, consideration must be given to whether it is a valid and reliable measure for that population of participants and raters. Care should be taken to report any steps taken, such as a pilot project or content expert review.

Examples in the literature in which these kind of activities are demonstrated include a study by Reinhardt et al. (2012) in which the LCJR was adapted from its original purpose (evaluating clinical judgment) for evaluation of a clinical skill. Another example used instruments originally designed for undergraduate nursing students to evaluate an interprofessional simulation experience among already licensed health care professionals (Prentice et al., 2011). Although the report details the instruments' original reliability, it does not discuss how the Simulation Design Scale, the Educational Practices in Simulation Scale, and the Self-Confidence in Learning Scale were evaluated for use with this new population.

Finally, researchers should consider their options for evaluation in light of the TSR and Kirkpatrick frameworks for categorizing simulation evaluation. The literature is saturated with reports of low-level participant evaluations, including reaction (Kirkpatrick's Level 1). It is time to step up and focus on what really matters: how simulation affects learning, behaviors, and ultimately patient outcomes.

Conclusion

Researchers can assist the continued maturation of the simulation pedagogy by aspiring to higher levels of evaluation and reporting psychometric measures and steps taken to assure validation with new populations. This report included instruments developed in several countries. Sharing the results of study replication from different cultural and international environments is an essential part of the further development of valid and reliable measures for simulation instruments. Replication studies using existing instruments with new populations and venues will be part of the process to turn tentative belief into accepted knowledge. Replications help further establish reliability, validity, and practice (Haller & Reynolds, 1986).

References

Ackermann, A. D. (2009). Investigation of learning outcomes for the acquisition and retention of CPR knowledge and skills learned with the use of high-fidelity simulation. *Clinical Simulation in Nursing*, *5*(6), e213-e222. http://dx.doi.org/10.1016/j.ecns.2009.05.002.

Adamson, K. A. (2012). Piloting a method for comparing two experiential teaching strategies. *Clinical Simulation in Nursing*, *8*(8) e375-e382. doi: 10.1016/j.ecns.2011.03.005.

Adamson, K. A., Gubrud, P., Sideras, S., & Lasater, K. (2012). Assessing the reliability, validity, and use of the Lasater Clinical Judgment Rubric: Three approaches. *Journal of Nursing Education*, *51*(2), 66-73. http://dx.doi.org/10.3928/01484834-20111130-03.

Adamson, K. A., Parsons, M. E., Hawkins, K., Manz, J. A., Todd, M., & Hercinger, M. (2011). Reliability and internal consistency findings from the C-SEI. *Journal of Nursing Education*, *50*(10), 583-586. http://dx.doi.org/10.3928/01484834-20110715-02.

Aronson, B., Glynn, B., & Squires, T. (in press-a). Competency assessment in simulated response to rescue events. *Clinical Simulation in Nursing*. doi: 10.1016/j.ecns.2010.11.006.

Aronson, B., Glynn, B., & Squires, T. (in press-b). Effectiveness of role-modeling intervention on student nurse simulation competency. *Clinical Simulation in Nursing*. doi: 10.1016.j.ecns.2011.11.005.

Boulet, J. R., Jeffries, P. R., Hatala, R. A., Korndorffer, J. R. J., Feinstein, D. M., & Roche, J. P. (2011). Research regarding methods of assessing learning outcomes. *Simulation in Healthcare*, *6*(7), e48-e51. http://dx.doi.org/10.1097/SIH.0b013e31822237d0.

Bray, B. S., Schwartz, C. R., Odegard, P. S., Hammer, D. P., & Seybert, A. L. (2011). Assessment of human patient simulation-based learning. *American Journal of Pharmaceutical Education*, *75*(10). http://dx.doi.org/10.5688/ajpe7510208. Article 208.

Brett-Fleegler, M., Monuteaux, M., Fleegler, E., Rudolph, J., Simon, R., Eppich, W., & Cheng, A. (October 01, 2012). Debriefing assessment for simulation in healthcare: Development and psychometric properties. *Simulation in Healthcare, 7*, 5, 288-294.

Brown, D., & Chronister, C. (2009). The effect of simulation learning on critical thinking and self-confidence when incorporated into an electrocardiogram nursing course. *Clinical Simulation in Nursing*, *5*(1), e45-e52. http://dx.doi.org/10.1016/j.ecns.2008.11.001.

Cason, C. L., & Baxley, S. M. (2011). Learning CPR with the Anytime™ for Healthcare Providers kit. *Clinical Simulation in Nursing*, *7*(6), e237-e243. http://dx.doi.org/10.1016/j.ecns.2010.06.002.

Cazzell, M., & Howe, C. (2012). Using objective structured clinical evaluation for simulation evaluation: Checklist considerations for interrater reliability. *Clinical Simulation in Nursing*, *8*(6), e219-e225. http://dx.doi.org/10.1016/j.ecns.2011.10.004.

Clark, M. (2006). Evaluating an obstetric trauma scenario. *Clinical Simulation in Nursing*, *2*(2), e75-e77. http://dx.doi.org/10.1016/j.ecns.2009.05.028.

Cohen, E. R., Feinglass, J., Barsuk, J. H., Barnard, C., O'Donnell, A., McGaghie, W. C., et al. (2010). Cost savings from reduced catheter-related bloodstream infection after simulation-based education for residents in a medical intensive care unit. *Simulation in Healthcare*, *5*(2), 98-102. http://dx.doi.org/10.1097/SIH.0b013e3181bc8304.

Cohen, S. R., Walker, D. M., & Marquez, F. E. (2010). PRONT02: Development, implementation and evaluation of an in-situ multi-disciplinary, low-tech, high-fidelity obstetric simulation training for Mexico. *Clinical Simulation in Nursing*, *6*(3), e110. http://dx.doi.org/10.1016/j.ecns.2010.03.017.

Davis, A. H., & Kimble, L. P. (2011). Human patient simulation evaluation rubrics for nursing education: Measuring the Essentials of Baccalaureate Education for Professional Nursing Practice. *Journal of Nursing Education*, *50*(11), 605-611. http://dx.doi.org/10.3928/01484834-20110715-01.

DeVita, M. A. (2009). Society for Simulation in Healthcare Presidential Address. *Simulation in Healthcare*, *4*(1), 43-48. http://dx.doi.org/10.1019/SIH.0b13e318197d315.

Donaghue, A., Nishisaki, A., Sutton, R., Hales, R., & Boulet, J. (2010). Reliability and validity of a scoring instrument for clinical performance during pediatric advanced life support simulation scenarios. *Resuscitation*, *81*, 331-336. http://dx.doi.org/10.1016/j.resuscitation.2009.11.011.

Dougherty, D., & Conway, P. H. (2008). The ''3t's'' road map to transform US health care: The ''how'' of high-quality care. *Journal of the American Medical Association*, *299*(19), 2319-2321. http://dx.doi.org/10.1001/jama.299.19.2319.

Dreifuerst, K. T. (2012). Using debriefing for meaningful learning to foster development of clinical reasoning in simulation. *Journal of Nursing Education*, *51*(4), 1-8. http://dx.doi.org/10.3928/01484834-20120409-02.

Elfrink Cordi, V. L., Leighton, K., Ryan-Wenger, N., Doyle, T. J., & Ravert, P. (2012). History and development of the Simulation Effectiveness Tool (SET). *Clinical Simulation in Nursing*, *8*(6), e199-e210. http://dx.doi.org/10.1016/j.ecns.2011.12.001.

Fero, L. J., O'Donnell, J. M., Zullo, T. G., Devito Debbs, A., Kitutu, J., Samosky, J. T., et al. (2010). Critical thinking skills in nursing students: comparison of simulation-based performance with metrics. *Journal of Advanced Nursing*, *66*(10), 2182-2193. doi:10.1111.j.1365-2646.2010.05385.x.

Fluharty, L., Hayes, A. S., Milgrom, L., Malarney, K., Smith, D., Reklau, M. A., et al. (2012). A Multisite, multi-academic track evaluation of end-of-life simulation for nursing education. *Clinical Simulation in Nursing*, *8*(4), e135-e143. http://dx.doi.org/10.1016/j.ecns.2010.08.003.

Gantt, L. T. (2010). Using the Clark Simulation Evaluation Rubric with associate degree and baccalaureate nursing students. *Nursing Education Perspectives*, *31*(2), 101.

Gore, T., Hunt, C. W., Parker, F., & Raines, K. H. (2011). The effects of simulated clinical experiences on anxiety: Nursing students' perspectives. *Clinical Simulation in Nursing*, *7*(5), e175-e180. http://dx.doi.org/10.1016/j.ecns.2010.02.001.

Grant, J., Epps, C., Moss, J., & Watts, P. (2009). Promoting reflective learning of students using human patient simulators. *Clinical Simulation in Nursing*, *5*(3), S6. http://dx.doi.org/10.1016/j.ecns.2009.03.172.

Grant, J. S., Moss, J., Epps, C., & Watts, P. (2010). Using video-facilitated feedback to improve student performance following high-fidelity simulation. *Clinical Simulation in Nursing*, *6*(5), e177-e184. http://dx.doi.org/10.1016/j.ecns.2009.09.001.

Guimond, M. E., & Simonelli, M. C. (2012). Development of the obstetric nursing self-efficacy Scale instrument. *Clinical Simulation in Nursing*, *8*(6), e227-e232. http://dx.doi.org/10.1016/j.ecns.2011.01.007.

Haller, K. B., & Reynolds, M. A. (1986). Using research in practice: A case for replication in nursing—Part Two. *Journal of Western Research*, *8*(2), 249-252.

Hayden, J. (2011). A national survey of simulation prevalence and use as clinical replacement: Results from Phase 1 of the NCSBN simulation study. *Clinical Simulation in Nursing*, *7*(6), e245. http://dx.doi.org/10.1016/j.ecns.2011.09.036.

Hemman, E. A., Gillingham, D., Allison, N., & Adams, R. (2007). Evaluation of the combat medic skills test. *Military Medicine*, *172*(8), 843-851.

Hutton, M., Coben, D., Hall, C., Rowe, D., Sabin, M., Weeks, K., et al. (2010). Numeracy for nursing, report of a pilot study to compare outcomes of two practical simulation tools—An online medication dosage assessment and practical assessment in the style of objective structured clinical examination. *Nurse Education Today*, *30*, 608-614. http://dx.doi.org/10.1016/j.nedt.2009.12.009.

Jeffries, P., & Rizzolo, M. (2006). Designing and implementing models for the innovative use of simulation to teach nursing care of ill adults and children: A national multi-site study. Retrieved October 8, 2012, from. http://www.nln.org/researchgrants/LaerdalReport.pdf.

Kardong-Edgren, S., & Adamson, K. A. (2009). BSN medical-surgical student ability to perform CPR in a simulation: Recommendations and implications. *Clinical Simulation in Nursing*, *5*(2), e79-e83. http://dx.doi.org/10.1016/j.ecns.2009.01.006.

Kardong-Edgren, S., Adamson, K., & Fitzgerald, C. (2010). A review of currently published evaluation instruments for human patient simulation. *Clinical Simulation in Nursing*, *6*(1), e25-e35. http://dx.doi.org/10.1016/jecns.2009.08.004.

Kirkpatrick, D. L. (1994). *Evaluating training programs: The four levels*. San Francisco, CA: Bernett-Koehler.

Klakovich, M. D., & Dela Cruz, F. A. (2006). Validating the interpersonal communication assessment scale. *Journal of Professional Nursing*, *22*(1), 60-67. http://dx.doi.org/10.1016/j.profnurs.2005.12.005.

Kogan, J. R., Holmboe, E. S., & Hauer, K. E. (2009). Tools for direct observation and assessment of clinical skills of medical trainees. *Journal of the American Medical Association*, *302*(12), 1316-1326. http://dx.doi.org/10.1001/jama.2009.1365.

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

Levett-Jones, T., McCoy, M., Lapkin, S., Noble, D., Hoffman, K., Dempsey, J., et al. (2011). The development and psychometric testing of the Satisfaction with Simulation Experience Scale. *Nurse Education Today*, *31*, 705-710, doi:10.1016.j.nedt.2011.01.004.

Liaw, Y. S., Sherpbier, A., Rethans, J.-J., & Klainin-Yobas, P. (2012). Assessment of simulation learning outcomes: A comparison of knowldege and self-reported confidence with observed clinical performance. *Nurse Education Today*, *32*(6), e35-e39. http://dx.doi.org/10.1016/j.nedt.2011.10.006.

Manz, J. A., Hercinger, M., Todd, M., Hawkins, K., & Parsons, M. E. (in press). Improving consistency of assessment of student performance during simulated experiences. *Clinical Simulation in Nurisng*. doi: 10.1016/j.ecns.2012.02.007.

Mariani, B., Cantrell, M. A., Meakim, C., Prieto, P., & Dreifuerst, K. T. (in press). Structured debriefing and students' clinical judgement abilities in simulation. *Clinical Simulation in Nursing*. doi: 10.1016/j.ecns.2011.11.009.

Massey, V. H., & Warblow, N. A. (2009). Using a clinical simulation to assess program outcomes. In M. H. Oermann, & K. T. Heinrich (Eds.), *Review of nursing education: Strategies for teaching, assessment, and program planning* (pp. 95-105). New York, NY: Springer.

McGaghie, W. C., Draycott, T. J., Dunn, F. W., Lopez, C. M., & Stefanidis, D. (2011). Evaluating the impact of simulation on translational patient outcomes. *Simulation in Healthcare*, *6*(7), S42-S47. http://dx.doi.org/10.1097/SIH.0b013e318222fde9.

McWilliam, P., & Botwinski, C. (2010). Developing a successful nursing objective structed clinical examination. *Journal of Nursing Education*, *49*(1), 36-41. http://dx.doi.org/10.3928/01484834-20090915-01.

Meyer, M. N., Connors, H., Hou, Q., & Gajewskil, B. (2011). The effect of simulation on clinical practice. *Simulation in Healthcare*, *6*(5), 269-277. http://dx.doi.org/10.1097/SIH.0b013e318223a048.

Mitchell, M. L., Henderson, A., Groves, M., Dalton, M., & Nulty, D. (2009). The objective structured clinical examination (OSCE): Optimising its value in undergraduate nursing curriculum. *Nurse Education Today*, *29*, 398-404. http://dx.doi.org/10.1016/.nedt.2008.10.007.

National Institutes of Health. (2011). Translational research. Retrieved July 5, 2012 from http://commonfund.nih.gov/clinicalresearch/overview-translational.aspx.

National Institutes of Health. (2012). Clinical and translational science. Retrieved September 11, 2012, from http://www.ncats.nih.gov/research/cts/cts.html.

Prentice, D., Taplay, K., Horsley, E., Payeur-Grenier, S., & Delford, D. (2011). Interprofessional simulation: An effective training experience for health care professionals working in community hospitals. *Clinical Simulation in Nursing*, *7*(2), e61-e67. http://dx.doi.org/10.1016/j.ecns.2010.03.001.

Radhakrishnan, K., Roche, J., & Cunningham, H. (2007). Measuring clinical practice parameters with human patient simulation: A pilot study. *International Journal of Nursing Education Scholarship*, *4*(1). Article 8.

Ravert, P., Williams, M., & Fosbinder, D. M. (1997). The Interpersonal Competence Instrument for Nurses. *Western Journal of Nursing Research*, *19*(6), 781.

Reed, S. (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.

Reinhardt, A. C., Mullins, I. L., De Blieck, C., & Schultz, P. (2012). IV insertion simulation: Confidence, skill and performance. *Clinical Simulation in Nursing*, *8*(6), e157-e167. http://dx.doi.org/10.1016/j.ecns.2010.09.001.

Rizzolo, M. A., Oermann, M. H., Jeffries, P., & Kardong-Edgren, S. (2011). NLN project to explore use of simulation for high states assessment. *Clinical Simulation in Nursing*, *7*(6), e245-e246. http://dx.doi.org/10.1016/j.ecns.2011.09.036.

Roche, J., Schoen, D., & Kruzel, A. (in press). Human patient simulation versus written case studies in new graduate nurses in nursing orientation: A pilot study. *Clinical Simulation in Nursing*. doi: 10.1016/j.ecns.2012.01.004.

Selim, A. A., Ramadan, F. H., El-Gueneidy, M. M., & Gaafer, M. M. (2012). Using objective structured clinical examination (OSCE) in undergraduate psychiatrice nursing education: Is it reliable and valid? *Nurse Education Today*, *32*, 283-288. http://dx.doi.org/10.1016/j.nedt.2011.04.006.

Senette, L., O'Malley, M., & Hendrix, T. (in press). Passing the baton: Using simulation to develop student collaboration. *Clinical Simulation in Nursing*.

Sevdalis, N., Undre, S., Henry, J., Sydney, E., Koutantji, M., Darzi, A., et al. (2009). Development, initial reliability and validity testing of an observational tool for assessing technical skills of operating room nurses. *International Journal of Nursing Studies*, *46*, 1187-1193.

Shinnick, M. A., & Woo, M. A. (2012). The effect of human patient simulation on critical thinking and its predictors in prelicensure nursing students. *Nurse Education Today.*

Shinnick, M. A., Woo, M., & Evangelista, L. S. (2012). Predictors of knowledge gains using simulation in the education of prelicensure nursing students. *Journal of Professional Nursing*, *28*(1), 41-47.

Swenty, C. F., & Eggleston, B. M. (2011). The evaluation of simulation in a baccalaureate nursing program. *Clinical Simulation in Nursing*, *7*(5), e181-e187, doi:org/10.1016/j.ecns.2010.02.006.

Tanner, C. (2011). The critical state of measurement in nursing education research. *Journal of Nursing Education*, *50*(9), 491-493.

Todd, M., Manz, J., Hawkins, K., Parsons, M., & Hercinger, M. (2008). The development of a quantitative evaluation tool for simulation in nursing education. *International Journal of Nursing Education Scholarship*, *5*(1). Article 41.

Woolf, S. H. (2008). The meaning of translational research and why it matters. *Journal of the American Medical Association*, *299*(2), 211-213.

Yuan, H. B., Williams, B. A., Fang, J. B., & Ye, Q. H. (2012). A systematic review of selected evidence on improving knowledge and skills through high-fidelity simulation. *Nurse Education Today*, *3*(3), 294-298. http://dx.doi.org/10.1016/j.nedt.2011.07.010.