Journal of Professional Nursing 35 (2019) 461–466

Contents lists available at ScienceDirect

Journal of Professional Nursing

journal homepage: www.elsevier.com/locate/jpnu

Reliability and validity of the simulation learning effectiveness inventory☆

Teresa A. Bates*, Patricia C. Clark

Byrdine F. Lewis College of Nursing and Health Professions, Georgia State University, United States of America

ARTICLE INFO

Keywords:
Simulation
Learning outcomes
Reliability
Validity
Factor analysis
Psychometric analysis

ABSTRACT

Background: Instruments developed to measure simulation learning outcomes need evidence of their reliability and validity for rigorous research. The purpose of this paper is to report psychometric properties of the English version of the Simulation Learning Effectiveness Inventory.

Methods: Psychometric properties of the English version of the Simulation Learning Effectiveness Inventory, which included internal consistency reliability and construct validity with factor analysis, were examined in a sample of 132 undergraduate nursing students.

Results: Cronbach's alpha coefficients were > 0.70 for all subscales. There was evidence of convergent, discriminant, and known-group validity. The factor analysis resulted in some items being associated with different subscales than in the original Chinese version.

Conclusions: The English version of the Simulation Learning Effectiveness Inventory has evidence of reliability and validity. Additional psychometric studies may result in changes in some of the subscales.

Introduction

Research about learning using simulation for healthcare students has increased substantially in recent years. With the growth of this teaching strategy, instruments to measure the process and outcomes of simulation experiences have been developed; however, at times do not have extensive evidence of reliability and validity prior to their use. In simulation articles, one or more instruments are developed by authors, but often there is little reported reliability and validity except expert review (e.g. Ha & Lim, 2018; Kirkpatrick et al., 2018; Nagelkerk et al., 2014; Wang, Liang, Blazeck, & Greene, 2015). Systematic reviews of simulation in nursing education (Doolen et al., 2016) and practice (Rutherford-Hemming & Alfes, 2017) have identified the need for reliable and valid instruments for rigorous measurement.

To conduct rigorous research about simulation learning, measures need to have evidence of reliability and validity to best determine if simulated clinical experiences result in expected learning benefits. The purpose of this paper is to report reliability and construct validity including exploratory factor analysis (EFA) for the English version of the Simulation Learning Effectiveness Inventory (SLEI) developed and reported by Chen, Huang, Liao, and Liu (2015). They developed the SLEI to capture multiple dimensions of the simulation experience from the student's perspective. Chen and colleagues reported evidence of seven subscales in the SLEI measuring the underlying dimensions of preparation, process, and outcomes of simulation experiences. Although the SLEI has been used in a simulation study (Bates, Moore, Greene, & Cranford, 2018), to our knowledge, this is the first reported evaluation of the psychometrics of the English version of this instrument in a sample of undergraduate nursing students in the United States.

Methods

This is a secondary data analysis from a study to examine active and observer roles in a simulation activity (Bates et al., 2018). There were 132 undergraduate nursing students completing the study. The simulation activity, along with the selection of instruments used in the study, was guided by the INACSL Standards of Best Practice: Simulation SM (INACSL Standards Committee, 2016) and the NLN Jeffries Simulation Theory (Jeffries, 2016). The details of the simulation activity are reported elsewhere (Bates et al., 2018). The National League of Nursing (NLN)'s SimMan scenario featuring Maria Gonzales, a 46-year-old Hispanic patient with acute pancreatitis (National League for Nursing, 2013) was used for the simulation. This NLN scenario was the first high-fidelity simulation (HFS) experience for all students. Students in a sophomore level, health assessment and basic skills course and in a junior level, medical/surgical course participated in this HFS. At the beginning of prebriefing, demographic data were collected and a trait anxiety tool

☆ This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
* Corresponding author at: Byrdine F. Lewis College of Nursing and Health Professions, 140 Decatur Street, Atlanta, GA 30302, United States of America.
E-mail address: tbates@gsu.edu (T.A. Bates).

https://doi.org/10.1016/j.profnurs.2019.04.007
Received 19 September 2018; Received in revised form 17 April 2019; Accepted 22 April 2019

was administered. At the end of prebriefing, students were randomly assigned a role and then a state anxiety tool was administered. At the end of debriefing, the state anxiety tool was used again along with the SLEI and the Student Satisfaction and Self-Confidence in Learning Scale (National League for Nursing, 2005).

Simulation Learning Effectiveness Inventory (SLEI)

The development of the SLEI was guided by the NLN Jeffries Simulation Framework and a thorough literature review performed by the scale developers (Chen et al., 2015). The results of the initial psychometrics of the SLEI that include EFA and confirmatory factor analysis (CFA) resulted in seven subscales: Course (3 items), Resource (4 items), Debriefing (4 items), Clinical Ability (5 items), Confidence (5 items), Collaboration (4 items), and Problem Solving (7 items). A second-order analysis showed that these subscales can be grouped into three major categories: Preparation, Process, and Outcome, corresponding with the components of simulation education. The component group scores are composed of the following subscales: Preparation (Courses, Resources), Process (Debriefing), and Outcome (Clinical Ability, Confidence, Problem Solving, Collaboration). The 32-item instrument has a 1–5 Likert-type response scale and the subscales can be scored individually or as three component group scores (Preparation, Process, and Outcome). Higher scores on the individual subscales or component group subscales indicate higher learning effectiveness of that domain (Chen et al., 2015).

In the initial development of the SLEI items, content validity was achieved through expert review, although there was no Content Validity Index (Waltz, Strickland, & Lenz, 2010) reported. Chen et al. used a large sample (N = 505) to conduct EFA and CFA of the SLEI. They also demonstrated discriminant and convergent validity of the instrument. The internal consistency reliability reported for the total scale and subscales was adequate.

Although the instrument was initially examined in Chinese, the instrument was published in English (Chen et al., 2015). The principal investigator of this study (Bates et al., 2018) adapted some wording in the English version with permission and discussion with the lead developer of the instrument (personal communication with Shiah Lian Chen). In addition, three nurse educators with expertise in simulation reviewed the instrument with the revised wording and found the revisions acceptable. For example, in the three items of the Course subscale, the word “course” was replaced with “simulation” as the students in our study were evaluating one simulation activity and not an entire simulation course with multiple simulations. In other items, the scale developers had used terms such as “situational learning,” “situational simulation practice,” and “situational discussion” which they viewed equal to simulation. Where these terms were stated in an item stem, they were replaced with “simulation” or “simulation exercise” to clarify that it was the simulation experience being evaluated. Also in two items, “medical team” was used and in one this was replaced with “team” and the other “healthcare team” as appropriate.

Student Satisfaction and Self-Confidence in Learning Scale

To examine convergent validity, the 13-item Student Satisfaction and Self-Confidence in Learning Scale (National League for Nursing, 2005) was used. The student's satisfaction with the simulation learning activity (5 items) and the student's self-confidence in learning (8 items) can be scored separately and were used separately here. This instrument has a 1–5 Likert-type response scale. Total scores for each subscale are computed by summing the items with the total possible scores from 5 to 25 for the Satisfaction subscale and 8–40 for the Self-Confidence subscale. Higher scores on each subscale indicated more satisfaction and self-confidence, respectively. Cronbach's alpha was reported as 0.94 for the Satisfaction subscale and 0.87 for the Self-Confidence subscale (Jeffries & Rizzolo, 2006). In our study, the Satisfaction subscale had a reliability coefficient of 0.90 and the Self-Confidence subscale coefficient was 0.78.

State-Trait Anxiety Inventory for Adults (STAI)

The well-established State-Trait Anxiety Inventory for Adults (STAI) was used (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 2015) for known-group and discriminate validity. State anxiety, or how an individual feels in response to a particular experience or situation, and trait anxiety, or an individual's proneness to anxiety, are each 20 item Likert-type response scales. State anxiety is expected to change depending on the situation and thus, the scale may be administered at more than one time point to evaluate change in anxiety over time. Trait anxiety is relatively stable and the scale is usually administered only at one time point. The total possible scores for each scale range from 20 to 80 with higher scores indicating a higher anxiety level. A cut-off point of ≥ 40 on the State Anxiety subscale has been suggested as clinically significant for symptoms of anxiety (Julian, 2011). The STAI has evidence of validity and reliability (Spielberger et al., 2015) and in our study the state anxiety at pretest and post-test had reliability coefficients of 0.95 and 0.94, respectively.

Data analysis

The α level for significance was set at $p < .05$. Cronbach's α coefficient was used to assess internal consistency reliability with an α coefficient of ≥ 0.70 considered acceptable. The Kaiser Meyer-Olkin measure of sampling adequacy and the Bartlett's test of sphericity were used to show that the sample and data were adequate for factor analysis (Field, 2009). Pearson's correlation coefficients were used to evaluate convergent and discriminate validity and Independent t-test for known-group validity (Di Iorio, 2005). Students were categorized as either high situational anxiety (STAI scores of ≥ 40) or low situational anxiety to examine known-group validity of the SLEI Confidence subscale and the Outcome component scores for the SLEI score. For correlational analysis, the continuous anxiety scores were used to examine discriminant validity with the SLEI's other subscales. See Table 1 for definitions of different types of validity.

Construct validity was assessed using EFA principal components with Promax rotation with fixed factors set at seven. This approach was used since the initial work reported the original EFA and later CFA resulted in seven factors with three latent factors. In this study, items with factor loadings ≥ 0.40 were retained on that factor (Field, 2009). Factor loadings represent the relationship of the item with the factor; that is, the variance that the item shares with the factor (Di Iorio, 2005).

Results

Table 2 provides the sample means, standard deviations, and actual minimum and maximum scores for the SLEI and other study instruments. The reliability coefficients of subscale and component scales also are reported in Table 2.

Internal consistency reliability of the SLEI

The Cronbach α coefficients for all the SLEI subscales and component group subscales were adequate ($\alpha \geq 0.70$) and are listed in Table 2. In examining the SLEI subscales, the item-to-total correlations were adequate and ranged for the Course subscale 0.66 to 0.74, Resource 0.52 to 0.75, Debriefing 0.74 to 0.85, Clinical Ability 0.66 to 0.83, Confidence 0.72 to 0.89, Problem Solving 0.55 to 0.74, and Collaboration 0.67 to 0.75. The “α if deleted” indicated that removing items would not substantially improve the internal consistency of any of the subscales. There were three items in the Confidence subscale and one in the Clinical Ability and Debriefing subscales with high item-to-

Table 1
Definitions of different types of validity for instruments and use in this study.

Type of validity	Definition
Content validity	"…the extent to which content represents the content domain…" (Waltz et al., 2010, p. 165). This is often obtained by expert review and/or use of a Content Validity Index (Waltz et al., 2010).
Construct validity	"… the extent to which relationships among items included in the measure are consistent with the theory and concepts as operationally defined" (Waltz et al., 2010 p. 167).
Construct validity - exploratory factor analysis (EFA) and confirmatory factor analysis (CFA)	EFA is used with new instruments "…to reduce a set of observed variables (items) to a smaller set of variables that reflects the interrelationships among the observed variables" (Di Iorio, 2005, p. 238). Results can provide support that items or subscales developed reflect the intended concepts within the measure. CFA is similar to EFA, except the analysis is used to confirm pre-defined different factors within an instrument or based on previous EFA findings (Di Iorio, 2005).
Construct validity - known-groups	Two groups expected to differ on the construct of interest are compared to determine if they significantly differ on the construct (Di Iorio, 2005). If the groups differ significantly in the direction anticipated, this adds to the validity of the measure. Thus, the measure can discriminate between groups. Example: Students with high situational anxiety compared to students with low situational anxiety were expected to differ on the SLEI Confidence subscale and the SLEI Outcome component scale.
Construct validity - convergent validity	Different measures of the same construct should be related (Waltz et al., 2010). Correlations often are used to determine if there is a significant relationship between two measures of the same or similar concepts. Example: Student confidence measured by the Satisfaction and Self-Confidence in Learning subscales was expected to be positively related to the SLEI confidence subscale.
Construct validity - discriminant (or divergent) validity	Correlations are examined with the construct of interest and other variables that are dissimilar (Di Iorio, 2005). The measures should not be related; thus, the correlations are expected to be low. Example: Pre-simulation state anxiety and trait anxiety were not expected to be related to any of the SLEI subscales except the Confidence subscale.

Table 2
Descriptive statistics for the SLEI and scales, reliability coefficient and instruments for validity ($N = 132$).

Variable[a]	*M* (*SD*)	Actual min-max score	α
Simulation preparation	26.28 (3.42)	12–30	0.87
Course	13.18 (1.82)	3–15	0.84
Resource	17.41 (2.51)	10–20	0.81
Simulation process - debriefing	18.39 (2.15)	4–20	0.91
Simulation outcome	89.83 (11.84)	46–105	0.95
Clinical ability	22.36 (2.87)	13–25	0.88
Confidence	20.45 (4.00)	5–25	0.92
Problem solving	29.13 (4.12)	14–35	0.86
Collaboration	17.89 (2.28)	10–20	0.86
Student satisfaction & self confidence in learning	56.68 (6.50)	23–65	0.90
Satisfaction	22.54 (3.00)	6–25	0.92
Self-confidence	34.14 (3.98)	17–40	0.78
Anxiety			
Trait anxiety	38.45 (10.14)	21–69	0.93
Pre-simulation state anxiety	40.00 (12.95)	20–75	0.95

[a] Uses items from the original scale to compute these data.

total correlations (≥0.80) indicating possible redundancy of items and possibly an artificially high reliability coefficient.

Construct validity of the SLEI

Construct validity of the SLEI subscales and component group subscales was assessed through examining the relationships among the Student Satisfaction and Self-Confidence in Learning subscales (Table 3) and trait anxiety and pre-simulation state anxiety scores. The Satisfaction and Self-Confidence in Learning subscale scores were significantly associated with all of the SLEI subscales and the SLEI Preparation, Process, and Outcome component subscales. Higher satisfaction and self-confidence with simulation were associated with positive SLEI scores. The direction and strength of the correlations were in the expected direction and provide support for convergent validity of the SLEI and its components; however, the correlations are low enough (highest shared variance 59%) to support the SLEI subscales are measuring different concepts. For example, it is anticipated that higher satisfaction of students with the simulation would be associated with more positive views of the simulation course, resources, debriefing, problem solving, and other outcomes. Also, it is not expected that the correlation coefficients would be high ($r \geq 0.90$) since that may indicate that the items are measuring the same construct.

Table 3
Relationships of SLEI subscales and components with student satisfaction, confidence, and anxiety ($N = 132$).

Variable	Satisfaction	Self-confidence in learning	Trait anxiety[a]	State anxiety[a]
Course	0.77**	0.66**	−0.18*	−0.07
Resources	0.59**	0.54**	−0.16	−0.08
Debriefing[b]	0.69**	0.63**	−0.13	−0.14
Clinical ability	0.74**	0.67**	−0.15	−0.07
Problem solving	0.68**	0.68**	−0.18*	−0.04
Confidence	0.71**	0.70**	−0.19*	−0.20*
Collaboration	0.60**	0.60**	−0.07	−0.01
Preparation component	0.71**	0.62**	−0.19*	−0.07
Outcome component	0.77**	0.75**	−0.18*	−0.10

* $p < .05$.
** $p < .001$.
[a] Pre-simulation measures; Pearson's correlation coefficients.
[b] Debriefing scale is the only scale in Process component of the SLEI.

None of the SLEI subscales and component group subscales was significantly related to the pre-simulation state anxiety except for the expected positive relationship with the Confidence subscale. The lack of relationships with most scale components is evidence of discriminant validity. Specifically, one would not expect state (situational) anxiety to necessarily be related to how students viewed variables such as simulation resources, debriefing, and collaboration. For trait anxiety, there were no significant relationships with the Resources, Debriefing, Clinical Ability, and Collaboration subscales. There was a significant negative relationship of trait anxiety as expected with the Confidence subscale. The relationship of state and trait anxiety with the Confidence subscale shows convergent validity. There also were significant negative relationships with the Course and Problem Solving subscales and the Preparation component subscale. However, these four correlations (Table 3) were low in magnitude and with the highest shared variance of these 3% between trait anxiety and the Preparation component subscale indicating little overlap between the scales. Therefore, there is evidence of discriminate validity for the SLEI subscales.

Construct validity through known-group assessment

We expected students with high state (situational) anxiety levels to have less positive learning outcomes (Holland, Gosselin, & Mulcahy, 2017; Mills, Carter, Rudd, Claxton, & O'Brien, 2016; Pai, Ko, Eng, & Yen, 2017). Students were categorized as either high state anxiety (STAI scores of ≥40) or low state anxiety to compare two groups on the SLEI Confidence subscale and the SLEI Outcome component subscales. We expected students with high state anxiety would have less confidence and worse outcomes in simulation. Independent *t*-tests were conducted to determine if the students with high anxiety ($n = 64$) differed from students with lower anxiety ($n = 68$) on the SLEI Confidence subscale and the SLEI Outcome component subscale. During the pre-simulation phase, high state anxiety students scored significantly lower ($M = 19.67 \pm 4.22$) on the SLEI Confidence subscale ($M = 21.19 \pm 3.65$; $t = 2.214$, $p = .029$). On the SLEI Outcome component subscale score (four combined SLEI individual subscales), there was no significant difference between those students with higher pre-simulation state anxiety ($M = 88.77 \pm 11.97$) compared to these with lower state anxiety ($M = 90.84 \pm 11.70$; $t = 1.006$; $p < .316$) on the overall simulation outcomes. These results provide partial support for known-group validity of the confidence learning outcome of the SLEI.

Construct validity through factor analysis

An EFA was conducted for the SLEI to determine the factor structure of the scale items. EFA using Principal Component analysis with Promax rotation forcing seven factors was conducted. The Kaiser-Meyer-Olkin for sampling adequacy was very high at 0.936 and the Bartlett's was statistically significant ($\chi^2 = 3448.056$; $p < .000$) indicating that the sample and data were adequate for factoring (Field, 2009). There was 16% of nonredundant residuals with absolute values of > 0.05 (smaller % is better), well below acceptable of 50% (Field, 2009). Six factors had Eigenvalues greater than one and the seventh factor had an Eigenvalue of 0.883. The seven factors accounted for a total of 75% of variance of the measure. The first factor contributed 49.8% of the total variance.

The items associated with the seven factors and their loadings on the primary factor are in Table 4 with the mean and standard deviation for each item. Some of the items loaded (factors loadings ≥ 0.40) on different factors compared to the original SLEI subscales (Chen et al., 2015). Generally, the Resource (Factor IV), Debriefing (Factor I), Confidence (Factor II), and Collaboration (Factor III) subscales had items that loaded as expected from the original scale except for a few differences. The original Course (Factor V), Problem Solving (VI), and Clinical Ability (Factor VII) subscales had items that loaded with different factors. All four of the original Debriefing subscale items loaded on Factor 1, but two additional items (#2 from the Course subscale and #7 from the Resource subscale; See Table 4 for items) loaded with the Debriefing items (Factor 1). The item (#2) from the Course subscale refers to knowing course objectives and evaluation requirements and the Resource subscale item (#7) referred to help always being available.

Factor II was the Confidence subscale items except for one item #17. Two additional items (#14 from the Clinical Ability subscale and #22 from the Problem Solving subscale) loaded on the Confidence subscale (Factor II). Item #14 was "This simulation contributed to my mastering the process of clinical care" and it also crossloaded on the Clinical Ability subscale (Factor VII), with an almost equal factor loading. However, the loading with the Confidence subscale (Factor II) was slightly higher and the Confidence subscale was a better conceptual fit for this item (Table 4).

Factor III was the Collaboration subscale items with one additional item (#28, Table 4) from the Problem Solving subscale loading on this factor. Item #28 addresses problem solving in confronting patient problems. Factor IV was the Resource subscale with all items except one. Item #7 from the Resource subscale (Table 4) loaded with the Debriefing items (Factor I).

Factor V included two items from the Course subscale (#1 & 3), one item was from the Clinical Ability subscale (#12), and the other one from the Problem Solving subscale (#23). Factor VI had three items (#24, 25, & 26) from the original Problem Solving subscale and Factor VII had two items (#15 & 16) from the original Clinical Ability subscale. As mentioned earlier, Item # 14 that addressed mastering the processes of clinical care crossloaded on Factor II with Confidence subscale items. There were two items (#13 & 17) that did not load on any factor, with one from the original Clinical Ability subscale and the other from the Confidence subscale.

The correlations of the seven factors were all positive. Most of the correlation coefficients among the factors were moderate, ranging from 0.22 to 0.65 with the exception of the correlation of Factor VI (Problem Solving) with Factor V (Course) and VII (Clinical Ability) which were 0.015 and 0.032, respectively. These correlations indicate most of the factors are related, but low enough so that the factors are not redundant. The correlations between Problem Solving (Factor V) and Clinical Ability (Factor VII) were very low indicating no relationship between these subscales.

We computed Cronbach alpha reliability coefficients for the seven factors with the items as they loaded on the Factors I-VII in Table 4. The alpha reliability coefficient was 0.89 for Factor I (Debriefing - 6 items), 0.93 for Factor II (Confidence - 7 items), 0.88 for Factor III (Collaboration - 5 items), 0.82 for Factor IV (Resources - 3 items), 0.83 for Factor V (Course - 4 items), 0.74 for Factor VI (Problem Solving - 3 items) and 0.71 Factor VII (Clinical Ability - 2 items). For all the items in each of the subscales based on the factors from the EFA, the item-to-total correlations were > 0.50.

Discussion

Nursing simulation experts recommend the reuse, psychometric reassessment, and revision of existing evaluation instruments (Adamson, Kardong-Edgren, & Willhaus, 2013; Kardong-Edgren, Adamson, & Fitzgerald, 2010). Building on the reliability and validity of currently published tools, through their use with varying student groups in multiple settings, may be a more efficient method of advancing the science of simulation as opposed to focusing on instrument development. The SLEI was chosen as an appropriate tool, instead of developing a new one, due to its initial reports of reliability and validity and since it was based on the NLN Jeffries Simulation Framework, a precursor of the theory that guided the original study (Bates et al., 2018). We found students in the observer role of a simulation experience had similar learning outcomes as students in an active simulation role using the SLEI (Bates et al., 2018).

The subscales of the English version of the SLEI used in this study have preliminary evidence of adequate internal consistency reliability; however, our factor analysis results differed somewhat from the initial results reported by Chen et al. (2015). The SLEI was initially conceptualized and had empirical support for seven first order factors and three underlying second order factors. Although we had the same subscale constructs, some of the items loaded on different subscales in our analysis. The Clinical Ability and Problem Solving subscales' items were particularly problematic.

Our findings highlight the need for evaluating the psychometric data of new instruments and particularly those that are translated from different languages. There was some evidence of convergent, discriminant, and known-group validity. There were some items with high item-to-total correlations indicating possible redundancy of these items. Additional research is needed to better determine if these items should be dropped or the wording revised and reexamined.

Table 4
Factors identified using exploratory factor analysis with rotation.

Original categories of items	Primary loading	Item M ± SD
Factor I - debriefing		
2. I understand the objective and evaluation requirements of this *simulation*. (originally on Course subscale)	0.564	4.47 ± 0.66
7. If I experienced problems or difficulty using the equipment, help was always available. (originally on Resource subscale)	0.475	4.31 ± 0.81
8. The teacher provided appropriate positive feedback according to the learning situation of students.	1.115	4.64 ± 0.60
9. The feedback provided by the teacher was immediate and promoted my learning outcome.	1.027	4.60 ± 0.63
10. Discussion with the teacher *during debriefing* assisted my achieving the learning goals.	0.894	4.58 ± 0.61
11. Feedback and discussion of the simulation assisted me in correcting my mistakes and promoting my learning.	0.790	4.58 ± 0.61
Factor II - confidence		
14. *This simulation* contributed to my mastering the processes of clinical care. (originally with Clinical Ability subscale)	0.423*	4.36 ± 0.82
18. *This simulation* practice boosted my confidence in my clinical skills.	0.663	4.11 ± 0.91
19. *This simulation exercise* boosted my confidence in handling future clinical problems.	0.658	4.12 ± 0.91
20. *This simulation exercise* alleviated my anxiety/fear of confronting future clinical patient problems.	1.018	3.72 ± 1.06
21. *This simulation exercise* contributed to my confidence in future patient care.	0.843	4.06 ± 0.86
22. *This simulation exercise* enabled me to understand the implication of each solution to patient problems.	0.898	4.16 ± 0.81
27. In participating in *this simulation*, I identified solutions to *problems by assessing and thinking about options*. (Originally Problem Solving subscale)	0.465	4.29 ± 0.64
Factor III - collaboration		
28. *This simulation exercise* promoted my problem-solving skills in confronting patient problems.	0.523	4.33 ± 0.69
29. *This simulation exercise* provided opportunities to practice communicating and cooperating with other members in my team.	0.894	4.53 ± 0.66
30. *This simulation* enabled me to understand the role that I should play in an interaction with a *healthcare* team.	0.711	4.44 ± 0.68
31. During the interaction in the *simulation*, I was willing to share *the* workload with other team members.	0.878	4.48 ± 0.66
32. I could discuss patient needs with the team by using effective communication skills.	0.719	4.44 ± 0.71
Factor IV - resources		
4. The equipment and resources for this *simulation* were sufficient.	0.847	4.14 ± 0.93
5. The equipment and resources for this *simulation* contributed to my learning.	0.712	4.47 ± 0.68
6. Using the environment and equipment for *simulation* was convenient.	0.885	4.48 ± 0.69
Factor V - course		
1. The *simulation* was arranged adequately in terms of sequential order and depth, facilitating my learning.	0.441	4.30 ± 0.75
3. The activities in this *simulation* assisted my achieving the learning goals.	0.531	4.41 ± 0.68
12. This *simulation* enhanced my understanding of patient problems. (Originally on Clinical Ability subscale)	0.668	4.57 ± 0.63
23. *This simulation exercise* enabled me to identify problems in clinical care that I have not noticed before. (Originally on Problem Solving subscale)	0.868	4.23 ± 0.82
Factor VI - problem solving		
24. In participating in *this simulation*, I approached new concepts or ideas through observation.	0.682	4.30 ± 0.69
25. *This simulation* enabled me to learn previously unfamiliar learning methods.	0.553	4.02 ± 0.88
26. In participating in *this simulation*, I approached solutions to problems *through looking up information I did not know*.	0.642	3.80 ± 1.02
Factor VII – clinical ability		
15. *This simulation* enabled me to acquire useful knowledge about clinical practices.	0.583	4.46 ± 0.66
16. The contents of *this simulation* corresponded with my previous learning experience.	0.669	4.48 ± 0.64
Items that did not load on any factor		
13. *This simulation* promoted my ability to care for patients. (Clinical Ability subscale)	–	4.48 ± 0.72
17. *This simulation exercise* encouraged me to confront future clinical challenges. (Confidence subscale)	–	4.44 ± 0.83

Item #14 which originally was on the Clinical Ability subscale crossloaded on Factor VII with 0.408. Italicized words indicate changes used in this study from original English version.

In assessing student outcomes of simulation, educators and researchers need to evaluate the different components of the learning experience, including the context, background, simulation experience, facilitator, educational strategies, and participants (Jeffries, 2016). The SLEI was constructed to assess the student participant's perceived learning effectiveness of one or more simulation experiences within a nursing course (Chen et al., 2015). The student participant evaluates components related to the context and background for the simulation, which includes the simulation design, goals and objectives, and sufficient equipment, time and facilitators. A simulation experience with good content may be ineffective if poorly organized or conducted by inexperienced educators. The actual, simulated learning activity, debriefing process, and educational strategies employed by the facilitators are the other main variables that need evaluation. It is crucial to determine whether the simulated client scenario provides the realism and opportunities for practicing skills needed to achieve the desired outcomes. Then the debriefing process is used as a time period for students to reflect on their actions, attitudes, and knowledge in order to improve future performance, communication, and decision making. Student participant learning outcomes are not only achieved during the simulation exercise, but also during the debriefing phase of simulation (INACSL Standards Committee, 2016).

Identifying clear definitions for the relevant concepts of simulation-based learning (e.g. clinical ability, problem solving, confidence) and prioritizing those that are most critical for improving the delivery of client care are essential for evaluation. For example, while students' perceived confidence in their clinical ability may be important, students can have “false” confidence based on feelings, but not be competent when held to a standard. It may be that exhibiting competence during a clinical simulation exercise or the next time the student is in the clinical area is the more crucial attribute or learning outcome. Measuring competence is likely better performed through observation of the skill instead of self-report. However, this method of measurement often is more time consuming. Thus, development of self-report scales that have evidence of being significantly related to more objective methods of measurement may be needed.

Generally, the SLEI item means were high as were the subscale scores indicating there may be a ceiling effect. This may limit the opportunity to detect change in simulation learning scores over time. Rewording items and conducting additional validity testing may help increase variability.

A few of the SLEI items as written do not follow some of the principals of writing questionnaire items (Di Iorio, 2005), such as having items include only one idea and being specific. Examples from Table 3 include (Item 2) "I understand the objective and evaluation requirements of this simulation" and (Item 29) "This simulation exercise provided me with opportunities to practice communicating and cooperating with other members of my team." These items include two ideas "objective and evaluation requirements" and "practice communicating and cooperating." These may need to be separate items. Also, two items (# 4 and 5), include both "equipment and resources" which may need to be separated in the item stems for clarity.

In one of the items (# 17, Table 3) that did not load on a factor, the wording seems problematic. The item "This simulation exercise encouraged me to confront future clinical challenges" may be considered an anthropomorphism as a simulation cannot "encourage." Thus, additional revisions to item stems may help with clarification being conceptually congruent with the proposed factor. In addition, once a researcher makes revisions to items, obtaining a review with content and measurement experts to obtain a Content Validity Index and qualitative feedback may be beneficial (Di Iorio, 2005).

This study has limitations and strengths. Although the sample size was adequate for the analysis conducted, larger samples are needed to conduct EFA and CFA. The English version of the SLEI used, with minor wording changes, was the one reported in Chen et al. (2015) and in that article there are no details of the translation process used for the instrument.

Conclusions

Currently, researchers may want to use the English version of the SLEI with caution and report separate subscale scores until additional psychometric testing better elucidates the conceptual components measured. However, the SLEI does focus on all aspects of the simulation process from planning to debriefing to learning outcomes, as guided by the NLN Jeffries Simulation Framework. Although in simulation-based learning, the focus is on the interactive experience of the student with the manikin or standardized patient, evidence-based preparation for the simulation is necessary in order to conduct a successful experience. If the planning or debriefing of a simulation is unorganized or poorly executed, even the best simulated client scenario may not result in the expected learning outcomes.

Simulation experiences are valuable learning opportunities for students to practice clinical skills, increase evidence-based knowledge, examine pre-existing attitudes, and explore professional role development. Nurse educators in academic and practice settings need to be able to effectively evaluate simulation experiences to determine which ones are most effective in achieving learning outcomes. Thus, we need instruments with evidence of reliability and validity to conduct rigorous evaluation of simulation-based learning activities.

References

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

Bates, T. A., Moore, L. C., Greene, D., & Cranford, J. S. (2018). Comparing outcomes of active student and observer roles in nursing simulation. *Nurse Educator*. https://doi.org/10.1097/NNE.0000000000000603 (Advance online publication).

Chen, S. L., Huang, T. W., Liao, I. C., & Liu, C. (2015). Development and validation of the simulation learning effectiveness inventory. *Journal of Advanced Nursing, 71*, 2444–2453. https://doi.org/10.1111/jan.12707.

Di Iorio, C. K. (2005). *Measurement in health and behavior methods for research and evaluation.* San Francisco: Jossey-Bass.

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

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Los Angeles: Sage.

Ha, E., & Lim, E. J. (2018). Peer-led written debriefing versus instructor-led oral debriefing: Using multimode simulation. *Clinical Simulation in Nursing, 18*, 38–46. https://doi.org/10.1016/j.ecns.2018.02.002.

Holland, B., Gosselin, K., & Mulcahy, A. (2017). The effect of autogenic training on self-efficacy, anxiety, and performance on nursing student simulation. *Nursing Education Perspectives, 38*, 87–89. https://doi.org/10.1097/01.NEP.0000000000000110.

INACSL Standards Committee (2016). INACSL standards of best practice: Simulation[SM]. *Clinical Simulation in Nursing, 12*, S5–S50 (doi:0.1016/j.ecns.2016.10.001).

Jeffries, P. R. (Ed.). (2016). *The NLN Jeffries simulation theory*. Philadelphia: Wolters Kluwer.

Jeffries, P. R., & Rizzolo, M. A. (2006). NLN-Laerdal project summary report: Designing and implementing models for the innovative use of simulation to teach nursing care of ill adults and children: A national multisite study. Retrieved from http://www.nln.org/docs/default-source/professional-development-programs/read-the-nln-laerdal-project-summary-report-pdf.pdf?sfvrsn=0.

Julian, L. J. (2011). Measures of anxiety. *Arthritis Care & Research, 63*, S467–S472. https://doi.org/10.1002/acr.20561.

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

Kirkpatrick, A., Ball, S., Connelly, S., Hercinger, M., Hanks, J., Potthoff, M., ... McCafferty, K. (2018). Intraprofessional simulation's impact on advanced practice and baccalaureate student self-efficacy. *Clinical Simulation in Nursing, 16*, 33–39. https://doi.org/10.1016/j.ecns.2017.11.005.

Mills, B., Carter, O., Rudd, C., Claxton, L., & O'Brien, R. (2016). An experimental investigation into the extent social evaluation anxiety impairs performance in simulation-based learning environments amongst final-year undergraduate nursing students. *Nurse Education Today, 45*, 7–15. https://doi.org/10.1016/j.nedt.2016.06.006.

Nagelkerk, J., Peterson, T., Pawl, B. L., Terman, S., Anyangu, A. C., Mlynarczyk, S., & Baer, L. J. (2014). Patient safety culture transformation in a children's hospital: An interprofessional approach. *Journal of Interprofessional Care, 28*, 358–364. https://doi.org/10.3109/13561820.2014.885935.

National League for Nursing (2005). Student satisfaction and self-confidence in learning. Retrieved from http://www.nln.org/professional-development-programs/research/tools-and-instruments/descriptions-of-available-instruments.

National League for Nursing (2013). Simulation in nursing education volume II scenarios for SimMan. New York: Laerdal Medical Corporation.

Pai, H. C., Ko, H. L., Eng, C. J., & Yen, W. J. (2017). The mediating effect of self-reflection and learning effectiveness on clinical nursing performance in nursing students: A follow-up study. *Journal of Professional Nursing, 33*, 287–292. https://doi.org/10.1016/j.profnurs.2017.01.003.

Rutherford-Hemming, T., & Alfes, C. M. (2017). The use of hospital-based simulation in nursing education: A systematic review. *Clinical Simulation in Nursing, 13*, 78–89. https://doi.org/10.1016/j.ecns.2016.12.007.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (2015). State-trait anxiety inventory for Adults™ manual. Retrieved from http://www.mindgarden.com.

Waltz, C. F., Strickland, O. L., & Lenz, E. R. (2010). *Measurement in nursing and health research* (4th ed.). New York: Springer Publishing Company, Inc.

Wang, W., Liang, Z., Blazeck, A., & Greene, B. (2015). Improving Chinese nursing students' communication skills by utilizing video-stimulated recall and role-play case scenarios to introduce them to the SBAR technique. *Nurse Educator Today, 35*, 881–887. https://doi.org/10.1016/j.nedt.2015.02.010.