

## ORIGINAL RESEARCH

# Translation and validation of the traditional Chinese NLN educational practices questionnaire, simulation design scale and student satisfaction and self-confidence in learning

Baljit Kaur Gill\*

*The Open University of Hong Kong, Hong Kong, China*

**Received:** March 29, 2020

**Accepted:** May 9, 2020

**Online Published:** May 17, 2020

**DOI:** 10.5430/jnep.v10n8p47

**URL:** <https://doi.org/10.5430/jnep.v10n8p47>

## ABSTRACT

**Background and objective:** Globally, the use of clinical simulation has been incorporated in different nursing programs. It is important to evaluate simulation using reliable and valid instruments. Using the same instrument helps to evaluate simulation under the same criteria both nationally and internationally. The National League of Nursing developed three simulation scales which is widely used in different countries and demonstrates a good reliability and validity. Nevertheless, it is only available in English. The aim of the study was to translate the original NLN simulation evaluation scales into Traditional Chinese and evaluate its psychometric properties.

**Methods:** Beaton and colleague's (2000) cross-cultural adaptation guidelines was adopted. Cronbach's alpha coefficient ( $\alpha$ ) and Corrected item-total correlation was used to determine the internal reliability. Haccoun's single group technique was used to assess the equivalent of the scale in the original and the translated version. Lastly, Exploratory Factor Analysis (EFA) was used to determine the factor structure and Intra-Class Correlation Coefficient (ICC) to test the stability of translated scale.

**Results:** Nine simulation experts from Hong Kong, Mainland China, Singapore and Taiwan confirmed translation of the NLN scales (EPQ-C, SDS-C, SSCL-C). Cronbach's alpha of all subscales and overall scales were acceptable (0.72-0.89). The intra-language, inter-language and temporal inter-language cross correlations between the original and translated scales were correlated ( $p < 0.01$ ). ICC of the translated scales ranges from good to excellent (0.78-0.91). Lastly, EFA also demonstrated the items were theoretically coherent ( $\geq 0.40$ ) and have the same factor structure as the original English version.

**Conclusions:** Traditional Chinese NLN simulation evaluation scales demonstrated strong validity and reliability.

**Key Words:** Simulation, Evaluation, Scales, Traditional Chinese, Translation, Psychometric analysis

## 1. INTRODUCTION

One of the top priorities in nursing education, is to help nursing students to apply their theoretical knowledge obtained in the classroom to clinical settings, which had led to the development and widely use of simulation as a teaching pedagogy in different nursing curriculums and programs.<sup>[1,2]</sup> Simulation had been widely used in western countries and there is

an increased emphasizes on using simulation in Asia's nursing education.<sup>[3,4]</sup> Simulation mimics real clinical setting and gives nursing students an opportunity of experiencing learning in a safe environment which enhances their development of affective, cognitive and psychomotor skills.<sup>[1]</sup> The use of simulation had also been strongly recommended by the World Health Organization [WHO] which helps to en-

\*Correspondence: Baljit Kaur Gill; Email: [dr.baljitkaurgill@gmail.com](mailto:dr.baljitkaurgill@gmail.com); Address: The Open University of Hong Kong, Hong Kong, China.

hance patient safety.<sup>[5]</sup> With its development, stakeholders are evaluating the design, implementation and outcomes of simulation at the same time. In order to enhance the development of simulation science, reliable and valid simulation instruments are important. Nevertheless, most of the simulation evaluation scales were developed in English, which may not meet the culture and language of Chinese simulation learners in Asia.

## 2. LITERATURE REVIEW

In the last two decades, simulation have been extensively used in nursing education in different regions with continuous evaluation of its use. But there is a relative lack of reliable and valid evaluation on simulation assessment instruments, measuring the learning outcomes and effectiveness of the teaching and learning strategies, which may inhibit its development, progress and adoption in nursing education. The pioneers in nursing simulation education recommends that further development and use of published simulation instruments is important,<sup>[6]</sup> and that using the same instrument in various nursing programs in different regions will help to determine the reliability and validity of the instrument in multiple settings and can therefore contribute to a consistent evaluation process. Using the same instrument will also help simulation users to evaluate and compare the effectiveness of simulation under the same criteria, both nationally and internationally. Thus, there is a need to test and validate the instruments and determining their relevance to nursing curriculum and programs in different countries.<sup>[2]</sup>

In the last 20 years, simulation had been widely used in Asia health care education namely, Mainland China, Hong Kong, Taiwan and Singapore and simulation learning centres were established.<sup>[7]</sup> Simulation programs and tactics effectively used in western countries may not be suitable for Chinese simulation learners in Asia. The learning style of a learner is often culturally-based and learners from different culture demonstrated a different learning, thinking and behavior.<sup>[8]</sup> Traditional Chinese is the main dialect in Hong Kong, Taiwan and Singapore. In Mainland China, simplified Chinese is widely used and it is derived from Traditional Chinese and most of the Mainlanders also understand Traditional Chinese. Therefore, simulation evaluation scales in Traditional Chinese will meet the need of most Chinese simulation learners in Asia.

### 2.1 National league of nursing (NLN) simulation evaluation scales

A preliminary literature search was conducted to explore commonly used simulation instruments in nursing education. National League of Nursing (NLN) had designed and

developed various scales namely, Educational Practices Questionnaire (EPQ), Simulation Design Scale (SDS) and Student Satisfaction and Self-Confidence in Learning scale (SSCL) in assessing the application and effectiveness of using simulation.<sup>[9,10]</sup> The scale development was supported by The NLN/Jeffires Simulation Framework which outlines the major concepts that contribute to nursing simulation research and evidence-based nursing practice<sup>[11]</sup> and it is commonly used in numerous simulation research. The original English version of the simulation evaluation scales had been translated and validated in various languages, i.e. Norwegian<sup>[12]</sup> and Turkish.<sup>[13]</sup>

#### 2.1.1 Educational practices questionnaire (EPQ)

Educational Practices Questionnaire (student version) (EPQ) measures four educational practices namely, active learning, collaboration, diverse ways of learning and high expectations. EPQ differentiates its rating according to the “Presence” and “Importance” of specific practices according to educational practice. EPQ demonstrates a high reliability (Cronbach’s alpha,  $\alpha$ ) in both the presence ( $\alpha = 0.86$ ) and importance ( $\alpha = 0.91$ ) scale.<sup>[14]</sup>

#### 2.1.2 Simulation design scale (SDS)

Simulation Design Scale (SDS) measures five features that are present in the simulation design namely, objectives of the simulation, the information available for the simulation, support provided in the simulation, problem-solving tactics, feedback and fidelity in the simulation. SDS differentiates its rating according to the “presence” and “importance” of specific practices scale. SDS demonstrates a high reliability (Cronbach’s alpha,  $\alpha$ ) in both the presence ( $\alpha = 0.92$ ) and importance ( $\alpha = 0.96$ ) scale.<sup>[15]</sup>

#### 2.1.3 Student satisfaction and self-confidence in learning scale (SSCL)

The Student Satisfaction and Self-Confidence in Learning scale (SSCL) measures student’s satisfaction of using simulation in learning and improvement of self-confidence in learning with simulation. SSCL demonstrates a high reliability in both satisfaction ( $\alpha = 0.94$ ) and self-confidence ( $\alpha = 0.97$ ).<sup>[15]</sup>

### 2.2 Aim of the study

The aim of this study was to translate NLN simulation evaluation scales into Traditional Chinese and evaluate its psychometric properties.

## 3. METHODS

Ethical approval for this study was obtained from the university where data collection was carried out. Permission was granted by the NLN to translate and use the three simulation evaluation scales (EPQ, SDS and SSCL).

### 3.1 Cross-cultural adaptation process

Beaton and colleague's (2000) cross-cultural adaptation guidelines was adopted and it includes six steps as 1) Forward translation; 2) Synthesis of the Translations; 3) Back Translation; 4) Expert Committee; 5) Test of the Pre-final Version; and lastly 6) Submission of documentation to the Developers or Coordinating Committee for Appraisal of the Adaptation Process.<sup>[16]</sup>

### 3.2 Content validity

Content validity of a scale demonstrates how the items in a scale match the concept that is being measured and is important when a scale is being designed.<sup>[17]</sup> The author, Pamela Jeffries, of the original NLN simulation evaluation scales considered the scales have a high content validity.<sup>[18]</sup> The expert committee members and competent translators in the translation process of the scales, serves well in the respective culture of both the sources and the target language countries. A focus group interview was also conducted to explore views of the participants on the feasibility and understanding on all the translated scales.

### 3.3 Sample selection & data collection

Undergraduate nursing students were recruited in a local university in Hong Kong. The inclusion criteria of the participants in the study are as follow:

- Participants should be bilingual, both fluent in English and Chinese;
- Participant's first language should be Traditional Chinese;
- Participants should have experienced simulation teaching and learning in the previous one month when completing the simulation evaluation scales.

Potential participants were approached after their simulation practice in the university. An information sheet explaining the study was given to the participant before completing the consent form of the study. Participants, who agreed to join the study, were invited to complete hard copies of original English (EPQ, SDS and SSCL) and the translated Traditional Chinese (EPQ-C, SDS-C and SSCL-C) version of NLN simulation evaluation scales.

### 3.4 Same size calculation

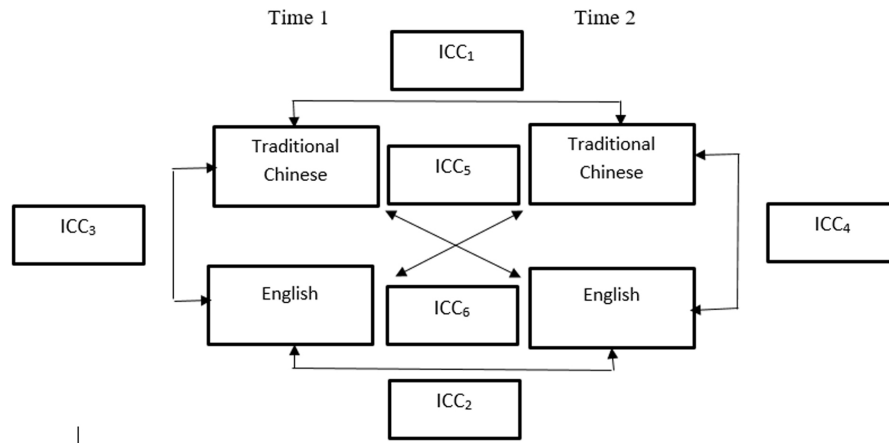
It is strongly recommended that there are at least 10 target population samples per item of the instrument for general psychometric approaches such as scale and item analysis, Pearson's Correlations and Exploratory Factor Analysis.<sup>[19-22]</sup> A minimum of 200 participants was required for performing psychometric analysis of the translated scales.

### 3.5 Statistical analysis

#### 3.5.1 Reliability and validity

Data analysis was conducted using the Statistical Package for the Social Sciences Statistical Package version 22.0 for Windows. Descriptive statistics were used to describe the characteristics of the recruited participants. Cronbach's alpha coefficient ( $\alpha$ ) and Corrected item-total correlation were used to determine the internal reliability and consistency of the translated scales. Cronbach's alpha with values above 0.70 was considered an acceptable level of reliability.<sup>[23]</sup> Corrected item-total correlation larger or equals to 0.4 was considered acceptable<sup>[24]</sup> and 0.4 was also used as a benchmark to determine the change of Cronbach's alpha if the item was removed.

Haccoun's single group technique was adopted and based on the assumption that the questionnaire was well translated, thus bilingual participants will provide equivalent responses to the question items in both the translated and original version of the scales at two different times.<sup>[25]</sup> The technique was administered at two different occasions, 4 weeks apart. It consists of a series of correlations that represent inter-language relationships and the index of the quality of the translation (see Figure 1). This method provides 1) test-retest intra-language correlations, measures the stability of the scale (English-English, Traditional Chinese-Traditional Chinese) at administration 1 and administration 2 ; 2) simultaneous inter-language correlations (English ↔ Traditional Chinese) indicate the quality of the translation is good and the scale provides the same information in both language and 3) temporal inter-language cross-correlations (English ↔ Traditional Chinese at administration 1 and 2) indicate that there is no interaction between time and original language of administration.<sup>[25]</sup> Pearson's correlation was used to measure the association between two variables.<sup>[26]</sup> Intraclass Correlations Coefficient (ICC) was used to measure intra language stability (test-retest reliability) at time 1 and time 2. Inter-language correlation was done at each time to determine whether the original and translated scales were consistent in both languages. Lastly, crossed correlation in the temporal inter-language was used to determine whether there was an interaction between the delay and the language of the initial administration. The inter-rater reliability coefficients can be interpreted as: 0.00 – 0.10 (virtually none); 0.11- 0.40 (slight); 0.41- 0.60 (fair); 0.61-0.80 (moderate) and 0.81-1.0 (substantial).<sup>[27]</sup> Intraclass Coefficient Correlations (ICC) was used to determine the stability of the translated scales. ICC of the translated scales was analyzed at four weeks apart and based on 95% confidence interval of ICC estimate, values less than 0.5, between 0.5- 0.75, between 0.75-0.9 and above 0.9 is interpreted as poor, moderate, good and excellent reliability respectively.<sup>[28]</sup>



**Figure 1.** Haccoun’s Unique Group Technique

ICC1 = Intra-language Chinese correlation (Chinese NLN 1st admin – Chinese 2nd admin); ICC2 = Intra-language Chinese correlation (English NLN 1st admin – English 2nd admin); ICC3 = Inter-language Chinese correlations (Chinese NLN 1st admin – English 1st admin); ICC4 = Inter-language Time 2 correlation (Chinese NLN 2nd admin – Chinese 2nd admin); ICC5 = Temporal Inter-language CIE2 correlation (Chinese NLN 1st admin – English 2nd admin); ICC6 = Temporal Inter-language C2E1 correlation (Chinese NLN 1st admin – English 2nd admin)

**3.5.2 Exploratory factor analysis (EFA)**

The translated simulation evaluation scales was translated for a new context, therefore exploratory factor analysis (EFA) was chosen which assumes no priori hypotheses about the dimensionality of a set of items.<sup>[29]</sup> Firstly, the data collected was tested for its eligibility before EFA using Kaiser-Meyer-Olkin (KMO) and Barlett’s test of sphericity. KMO value  $\geq 0.06$  demonstrates an adequacy of the sample for factor analysis<sup>[30]</sup> and  $p$  value smaller than .05 ( $p < .05$ ) in the Barlett’s test of Sphericity demonstrates that factor analysis is appropriate for the analysis.<sup>[30]</sup>

Principal Component Analysis (PCA) was adopted to obtain an empirical summary of the collected data source.<sup>[30]</sup> Principal factor method and varimax rotation were used to test whether all the items for a single factor loaded larger than 0.40 and it also confirms that the item loadings were theoretically coherent. Factors with item loading 0.40 and above were considered significant<sup>[29,31]</sup> and factors with eigenvalues over the point of inflexion were retained.<sup>[31]</sup>

**4. RESULTS**

**4.1 Sample characteristics**

355 eligible participants that meet the inclusion criteria were approached after their simulation session. 327 participants agreed to join the study and completed scales. Among the collected questionnaires, 27 of them, were found to have missing data in their collected questionnaires and were discarded. 300 completed questionnaires were included in the data analysis. The mean age of the students was 21.45 (SD 2.11) and students from four-year baccalaureate nursing pro-

gram was recruited (see Table 1).

**Table 1.** Demographic data of the participant (n = 300)

Characteristic of participant	n	%
Stream of Nursing Education		
General	247	82.30
Mental	53	17.70
Years of Undergraduate Nursing Education		
Year 1	70	23.30
Year 2	77	25.70
Year 3	75	25.00
Year 4	78	26.00

**4.2 Cross-adapted scale**

Two translators (T1 and T2) who use Traditional Chinese as their first language were invited to perform the forward translation from English to Traditional Chinese. One of the translators was with a nursing background and had experienced simulation teaching and learning in the past and the other was a high school teacher who specializes in incorporating technology in education. A finalized version of forward translation scale was developed (T12). Backward translation (BT1 & BT2) was performed by another two translators who do not have any knowledge of simulation. One of them was a Chinese medicine practitioner and the other is an administrative clerk. They were invited to translate the translated scale (T12) from traditional Chinese back to English (BT12).

**4.3 Content validity**

The consolidation of the scales was carried out by an advisory committee in term of semantic equivalence, idiomatic

equivalence, experiential equivalence and lastly conceptual equivalence. The discrepancies of words such as the Chinese name of “simulation” and layout of the likert scale with the rating sequence with the highest score either on the right or on the left of the likert scale were discussed. The finalized translated scales were reviewed by nine simulation experts from Hong Kong, China, Singapore and Taiwan and confirmed that the translated Traditional Chinese NLN simulation evaluation scales were clear and easy to understand, and demonstrate the same meaning as the original English version. Face-to-face focus group interview (n = 24) was also conducted by randomly selecting participants from the

study to explore their understanding and difficulties in completing the translated scales. Participants expressed that the translated scales were clear and easy to understand.

#### 4.4 Psychometric analysis

##### 4.4.1 Internal consistency

Cronbach’s alpha coefficient for presence and importance of specific practices is 0.87 and 0.91 for EPQ-C. The Cronbach’s alpha for presence and importance of specific practices of SDS-C is 0.91 and 0.95. SSCL-C only had two subscales, namely learner’s satisfaction and self-confidence. The Cronbach’s alpha coefficient for satisfaction and self-confidence is 0.85 and 0.78 for SSCL-C (see Table 2).

**Table 2.** Reliability (Cronbach’s alpha) of Translated Traditional Chinese and Original English version of EPQ-C, SDS-C and SSCL-C

		Presence of Specific Practices		Importance of Specific Practices	
<i>Subscale:</i>		<i>Chinese</i>	<i>English</i>	<i>Chinese</i>	<i>English</i>
<b>EPQ-C</b>	Active Learning (Item 1-10)	0.81	0.81	0.87	0.88
	Collaboration (Item 11-12)	0.73	0.74	0.80	0.82
	Diverse Ways of Learning (Item 13-14)	0.65	0.67	0.75	0.73
	High Expectations (Item 15-16)	0.72	0.75	0.76	0.77
	Whole Scale:	0.87	0.86	0.91	0.91
<i>Subscale:</i>		<i>Chinese</i>	<i>English</i>	<i>Chinese</i>	<i>English</i>
<b>SDS-C</b>	Objectives and Information (Item 1-5)	0.80	0.81	0.84	0.85
	Support (Item 6-9)	0.87	0.85	0.89	0.87
	Problem Solving (Item 10-14)	0.73	0.74	0.84	0.84
	Feedback/Guided Learning (Item 15-18)	0.80	0.82	0.86	0.85
	Fidelity (Realism) (Item 19-20)	0.724	0.74	0.87	0.87
Whole Scale:	0.91	0.92	0.95	0.96	
<i>Subscale:</i>		<i>Chinese</i>		<i>English</i>	
<b>SSCL-C</b>	Satisfaction with Current Learning (Item 1-5)	0.85		0.86	
	Self-Confidence in Learning (Item 6-13)	0.78		0.80	
	Whole Scale:	0.87		0.88	

Cronbach’s alpha of each subscale in both the presence and importance scales in EPQ-C, SDS-C and SSCL-C was also calculated (see table 2). Cronbach’s alpha of all subscale values above 0.70 except “Diverse Ways of Learning” subscale in presence scale of EPQ-C. In the translated NLN scales, all of the item values above 0.40 in corrected item-total correlation except item 2 (0.32) in the presence scale of EPQ-C (see Table 3) and item 13 (0.38) in the presence scale of SDS-C (see Table 4). All of the items in SSCL-S values above 0.40 (see Table 5).

Two-tailed *t*-tests and Pearson (*r*) product moment correlations were used to analysis the intra-language correlations,

inter-language correlations and temporal inter-language cross-correlations. Individual item in the scales was correlated for each analysis followed by subscale analysis (see Table 6). All correlations were found to be statistically significant at the 0.01 level.

##### 4.4.2 Test-retest reliability

Test and retest reliability of the translated scales was analyzed at four weeks interval. The Intraclass Coefficient Correlation (ICC) of the translated scales (EPQ-C, SDS-C and SSCL-C) based on 95% confidence ranges from good to excellent respectively (EPQ-C = 0.84, SDS-C = 0.91 and SSCL-C = 0.78).

**Table 3.** EPQ-C for presence and importance of specific practices with Corrected item-total correlation and Cronbach’s alpha if item deleted

	<i>Presence of Specific Practice</i>		<i>Importance of Specific Practice</i>	
	<b>Corrected item-total correlation</b>	<b>Cronbach’s alpha if item deleted</b>	<b>Corrected item-total correlation</b>	<b>Cronbach’s alpha if item deleted</b>
Q.1	0.48	0.86	0.63	0.91
Q.2	0.32	0.87	0.61	0.91
Q.3	0.44	0.86	0.53	0.91
Q.4	0.45	0.86	0.64	0.91
Q.5	0.59	0.86	0.67	0.91
Q.6	0.48	0.86	0.59	0.91
Q.7	0.56	0.86	0.60	0.91
Q.8	0.49	0.86	0.57	0.91
Q.9	0.57	0.86	0.57	0.91
Q.10	0.52	0.86	0.56	0.91
Q.11	0.50	0.86	0.60	0.91
Q.12	0.53	0.86	0.60	0.91
Q.13	0.46	0.86	0.49	0.91
Q.14	0.55	0.86	0.64	0.91
Q.15	0.54	0.86	0.60	0.91
Q.16	0.55	0.86	0.58	0.91

**Table 4.** SDS-C for presence and importance of specific practices with Corrected item-total correlation and Cronbach’s alpha if item deleted

	<i>Presence of Specific Practice</i>		<i>Importance of Specific Practice</i>	
	<b>Corrected item-total correlation</b>	<b>Cronbach’s alpha if item deleted</b>	<b>Corrected item-total correlation</b>	<b>Cronbach’s alpha if item deleted</b>
Q.1	0.59	0.90	0.62	0.94
Q.2	0.53	0.91	0.62	0.94
Q.3	0.60	0.90	0.70	0.94
Q.4	0.54	0.91	0.63	0.94
Q.5	0.58	0.91	0.61	0.94
Q.6	0.64	0.90	0.71	0.94
Q.7	0.59	0.90	0.68	0.94
Q.8	0.66	0.90	0.71	0.94
Q.9	0.67	0.90	0.76	0.94
Q.10	0.41	0.91	0.63	0.94
Q.11	0.43	0.91	0.63	0.94
Q.12	0.50	0.91	0.56	0.95
Q.13	0.38	0.91	0.66	0.94
Q.14	0.53	0.91	0.66	0.94
Q.15	0.61	0.90	0.69	0.94
Q.16	0.66	0.90	0.73	0.94
Q.17	0.56	0.91	0.65	0.94
Q.18	0.60	0.90	0.74	0.94
Q.19	0.42	0.91	0.68	0.94
Q.20	0.41	0.91	0.65	0.66

**Table 5.** SSCL-C Corrected item-total correlation and Cronbach’s alpha if item deleted

	Corrected item-total correlation	Cronbach’s alpha if item deleted
Q.1	0.66	0.82
Q.2	0.64	0.83
Q.3	0.65	0.83
Q.4	0.67	0.82
Q.5	0.72	0.80
Q.6	0.51	0.76
Q.7	0.47	0.76
Q.8	0.52	0.75
Q.9	0.57	0.75
Q.10	0.50	0.76
Q.11	0.40	0.77
Q.12	0.50	0.76
Q.13	0.43	0.77

**4.4.3 Exploratory factor analysis**

**EPQ-C**

Both presence ( $\chi^2 = 1456.05$ ;  $df = 120$ ;  $p \leq .001$ ) and importance ( $\chi^2 = 2076.87$ ;  $df = 120$ ;  $p \leq .001$ ) of specific practice in EPQ-C demonstrates significance of Bartlett’s test of sphericity. Kaiser-Meyer-Olkin (KMO) for both presence (KMO = 0.87) and importance (KMO = 0.91), adequate for factor analysis. An initial analysis using Principal Components Analysis (PCA) was run and revealed a four-factor structure both in the presence and importance scale of EPQ-C with eigenvalues above a Kaiser’s criterion of 1, thus explaining of 56.35% (33.99%, 8.77%, 7.23% and 6.36%) of total variance in EPQ-C presence and 62.32% (43.64%, 6.67%, 6.57% and 5.44%) in EPQ-C importance. The scree test also showed a clear break after the fourth component in both the presence and importance scale of EPQ-C. All factors in both the presence and importance scale rated above 0.40 with a same factor structure as the English version (see Table 7).

**Table 6.** Subscale intralanguage, interlanguage and temporal interlanguage cross-correlations for the English and Traditional Chinese version of the NLN simulation as assessment scales

NLN Simulation Assessment Scales	Pearson Correlation (r)					
	Intralanguage		Interlanguage		Cross-correlations	
	(E) (1)	(C) (1)	(E) (1)	(E) (2)	(E) (1)	(C) (1)
	Vs		Vs		Vs	
	(E) (2)	(C) (2)	(C) (1)	(C) (2)	(C) (2)	(E) (2)
<b><u>EPQ (Presence)</u></b>						
Active Learning	0.81	0.81**	0.85**	0.84**	0.85**	0.86**
Collaboration	0.72	0.74**	0.79**	0.80**	0.79**	0.79**
Diverse Ways of Learning	0.69	0.70**	0.72**	0.75**	0.78**	0.77**
High Expectations	0.74	0.73**	0.74**	0.76**	0.79**	0.78**
<b><u>EPQ (Importance)</u></b>						
Active Learning	0.84**	0.85**	0.84**	0.86**	0.86**	0.86**
Collaboration	0.80**	0.81**	0.82**	0.82**	0.84**	0.83**
Diverse Ways of Learning	0.78**	0.78**	0.79**	0.78**	0.80**	0.79**
High Expectations	0.78**	0.77**	0.77**	0.76**	0.78**	0.77**
<b><u>SDS (presence)</u></b>						
Objectives and Information	0.80**	0.81**	0.81**	0.81**	0.82**	0.81**
Support	0.84**	0.85**	0.86**	0.86**	0.87**	0.87**
Problem Solving	0.78**	0.77**	0.78**	0.78**	0.76**	0.77**
Feedback/Guided Reflection	0.80**	0.81**	0.81**	0.81**	0.82**	0.81**
Fidelity	0.73**	0.75**	0.73**	0.74**	0.74**	0.73**
<b><u>SDS (Importance)</u></b>						
Objectives and Information	0.84**	0.84**	0.85**	0.84**	0.86**	0.85**
Support	0.87**	0.89**	0.88**	0.89**	0.89**	0.89**
Problem Solving	0.84**	0.82**	0.84**	0.84**	0.82**	0.82**
Feedback/Guided Reflection	0.86**	0.87**	0.85**	0.86**	0.86**	0.85**
Fidelity	0.87**	0.87**	0.88**	0.87**	0.88**	0.89**
<b><u>SSCL</u></b>						
Satisfaction	0.85**	0.86**	0.86**	0.86**	0.87**	0.86**
Self-Confidence	0.79**	0.78**	0.79**	0.80**	0.80**	0.79**

\*\*Correlation is significant at the 0.01 level (2-tailed); E = Original English version; C = Translated Traditional Chinese version; 1 = 1<sup>st</sup> Administration; 2 = 2<sup>nd</sup> Administration

**Table 7.** Factor Loading of EPQ-C presence and importance of specific practices scale

	Presence of Specific Practice				Importance of Specific Practice			
	Subscale/Factors				Subscale/Factors			
	1	2	3	4	1	2	3	4
<b>Active learning</b>								
Q.1	0.79				0.79			
Q.2	0.69				0.79			
Q.3	0.62				0.49			
Q.4	0.54				0.48	0.37		
Q.5	0.53				0.89	0.39		
Q.6	0.51				0.76	0.36		
Q.7	0.79	0.69				0.86		
Q.8	0.75	0.39				0.82		
Q.9	0.66	0.31				0.55		
Q.10	0.84		0.33		0.32	0.53		
<b>Collaboration</b>								
Q.11	0.33	0.66				0.43		
Q.12		0.40				0.77		
<b>Diverse Ways of Learning</b>								
Q.13			0.83				0.74	
Q.14			0.68				0.71	
<b>High Expectations</b>								
Q.15				0.64				0.52
Q.16	0.36			0.47				0.40

Note. The shaded values indicate the best fit of items per subscale/factor (N = 300); 1 is subscale “Active Learning”, 2 is subscale “Collaboration”, 3 is subscale “Diverse Ways of Learning” and 4 is subscale “High Expectations”.

**SDS-C**

Both presence ( $\chi^2 = 2520.81$ ;  $df = 190$ ;  $p \leq .001$ ) and importance ( $\chi^2 = 3632.33$ ;  $df = 190$ ;  $p \leq .001$ ) of specific practice in SDS-C demonstrates significance of Bartlett’s test of sphericity. KMO for both presence (KMO = 0.91) and importance (KMO = 0.95) was adequate for factor analysis. The principal components analysis revealed a five-factor structure both in the presence and importance scale of SDS-C. SDS-C presence accounted 63.11% (37.78%, 8.41%, 6.30%, 5.95% and 4.67%) of the total variance and SDS-C importance accounted 69.65% (49.77%, 6.51%, 5.05%, 4.35% and 3.96%). The scree test also showed a clear break after the fifth component in both the presence and importance scale of SDS-C and all factors loaded above 0.40 with a same factor structure as the English version (see Table 8).

**SSCL-C**

SSCL-C demonstrated a significance of Bartlett’s test of sphericity ( $\chi^2 = 1350.06$ ;  $df = 78$ ;  $p \leq .001$ ) and KMO measures adequate samples for factor analysis in (KMO = 0.89). The principal components analysis revealed a two-factor structure and accounted 50.11% (40.38% and 9.79%) of the total variance. The scree test also showed a clear break

after the second component in SSCL-C and all the factors in the scale rated above 0.40 (see Table 9).

**5. DISCUSSION**

The strength of this study is that it had followed a stringent translation process with simulation experts from different countries in Asia. Bilingual participants were recruited from each year of study as they may have different levels of language competencies, simulation understanding and experience, collecting data from different years help to demonstrate that the translated scales can be used among simulation learners in different level of education. The majority of the participants were recruited from the general stream of nursing education, because a larger percentage of nursing students are enrolled in a general nursing stream in each year of the program which is also the percentage of employment in hospital settings in Hong Kong. Even though the participants were from two different streams, both of the streams were receiving the same simulation teaching and learning activities. Therefore, this confounding variable did not play any effect of the simulation experience and have any effect on the completion of the scales.



**Table 8.** Factor Loading of SDS-C presence and importance of specific practices scale

	Presence of Specific Practice					Importance of Specific Practice				
	Subscale/Factors					Subscale/Factors				
	1	2	3	4	5	1	2	3	4	5
<b>Objectives and Information</b>										
Q.1	0.83					0.88				
Q.2	0.80					0.78				
Q.3	0.80					0.72				
Q.4	0.77					0.64				
Q.5	0.65					0.70	0.50			
<b>Support</b>										
Q.6	0.57	0.60					0.70			
Q.7	0.46	0.51					0.66			
Q.8		0.79					0.40		0.34	
Q.9		0.64					0.76			
<b>Problem Solving</b>										
Q.10		0.59	0.61					0.75		
Q.11			0.42	0.36				0.71		
Q.12		0.52	0.82					0.64		
Q.13			0.79					0.37		
Q.14			0.78	0.58				0.89	0.59	
<b>Feedback/Guided Reflection</b>										
Q.15				0.78					0.85	
Q.16				0.64					0.81	
Q.17	0.35			0.45					0.74	
Q.18	0.39			0.43					0.71	0.69
<b>Fidelity (Realism)</b>										
Q.19					0.82					0.66
Q.20					0.82		0.46			0.83

Note. The shaded values indicate the best fit of items per subscale/factor (N = 300); 1 is subscale “Objectives and Information”, 2 is subscale “Support”, 3 is subscale “Problem Solving”, 4 is subscale “Feedback/Guided Reflection” and 5 is “Fidelity”.

**Table 9.** Factor Loading of SSCL-C

	Subscale/Factors	
	1	2
<b>Satisfaction with Current Learning</b>		
Q.1	0.87	
Q.2	0.84	
Q.3	0.81	
Q.4	0.78	
Q.5	0.67	
<b>Self-confidence in Learning</b>		
Q.6		0.55
Q.7	0.47	0.57
Q.8	0.44	0.51
Q.9	0.36	0.44
Q.10	0.30	0.43
Q.11		0.73
Q.12		0.70
Q.13		0.69

Note. The shaded values indicate the best fit of items per subscale/factor (N = 300); 1 is subscale “Satisfaction with Current Learning” and 2 is subscale “Self-Confidence in Learning”.

Subscale “Diverse Ways of Learning – item 13 and 14” in the presence scale of EPQ-C ( $\alpha = 0.65$ ) and EPQ ( $\alpha = 0.67$ ), demonstrated a slightly lower value of Cronbach’s alpha value both the English and Traditional Chinese. If item 13 and 14 are removed from the EPQ-C, the reliability still remains unchanged ( $\alpha = 0.86$ ). Therefore, the items will remain in the scale and considerations should be made on the results on this subscale. Corrected item-total correlations of item 2 (0.32) in the presence scale of EPQ-C is also slightly lower than 0.40, if the item is deleted in the scale, it only demonstrated a slight increase in Cronbach’s alpha value from 0.86 to 0.87. Hence, it remains in the translated scale EPQ-C. The Corrected item-total correlations of item 13 (0.38) of the presence scale of SDS-C is slightly lower than 0.40, if the item is deleted from the scale, the Cronbach’s alpha remains unchanged (0.91). Hence, it also remains in the translated SDS-C.

Haccoun’s single group technique was used to analyse intraglanuage correlations, interlanguage correlations and tem-

poral correlations analysed in each subscale and scale of the translated scales (EPQ-C, SDS-C and SSCL-C) demonstrated moderate to sufficient correlations (0.69 – 0.89). This confirms the reliability and validity of the translated scale is comparable to the English version. Intraclass Coefficients Correlations (ICC) at four weeks interval demonstrated a good to excellent reliability.

Factor analysis reviewed the factor structure of the translated scales (EPQ-C, SDS-C and SSCL-C). Factor loadings of each item in the scales were all larger than 0.40 which demonstrated that the items were theoretically relevant to the scale. The translated scales also demonstrated the same structure as the original English version.

### 5.1 Limitations

Participants were only limited to one region (Hong Kong) and participants from different regions in Asia is preferable. In the process of psychometric analysis of the original and translated NLN simulation evaluation scales, there was no gold standard of assessing EPQ, SDS and SSCL.

## 6. CONCLUSIONS

The use of simulation in nursing education is a global trend and it demonstrated a vast positive implication to the learners. In order to meet the demand of simulation in learning, different simulation programs were developed and shared among simulation educators as well. Simulation program's success in one nation may not be the same in the other culture. Therefore, it is important to evaluate the effectiveness of simulation among learners from different cultures. The translated Traditional Chinese NLN simulation evaluation scales (EPQ-C, SDS-C and SSCL-C) demonstrated strong reliability and validity which is equivalent to the original English version.

## ACKNOWLEDGEMENTS

The author, would like to thank Bobby Gill for his supports, encouragements, insights and guidance on doing this study. This work was supported by the Open University of Hong Kong Research Grant (No. 2014/1.1).

## CONFLICTS OF INTEREST DISCLOSURE

The author declares that there is no conflict of interest.

## REFERENCES

- [1] Cant RP, Cooper SJ. Simulation-based learning in nurse education: systematic review. *Journal of advanced nursing*. 2010; 66(1): 3-15. PMID:20423432 <https://doi.org/10.1111/j.1365-2648.2009.05240.x>
- [2] LaFond CM, Van Hulle Vincent C. A critique of the National League for Nursing/Jeffries simulation framework. *Journal of Advanced Nursing*. 2013; 69(2): 465-480. PMID:22715857 <https://doi.org/10.1111/j.1365-2648.2012.06048.x>
- [3] Wang AL, Fitzpatrick JJ, Petrini MA. Comparison of two simulation methods on Chinese BSN students' learning. *Clinical Simulation in Nursing*. 2013; 9(6): e207-e212. <https://doi.org/10.1016/j.ecns.2012.01.007>
- [4] Wang R, Shi N, Bai J, et al. Implementation and evaluation of an inter-professional simulation-based education program for undergraduate nursing students in operating room nursing education: a randomized controlled trial. *BMC Medical Education*. 2015; 15(1): 115. PMID:26155839 <https://doi.org/10.1186/s12909-015-0400-8>
- [5] World Health Organization: WHO patient safety curriculum guide: multi-professional edition. In.: WHO Geneva; 2011.
- [6] Kardong-Edgren S, Adamson KA, Fitzgerald C. A review of currently published evaluation instruments for human patient simulation. *Clinical Simulation in Nursing*. 2010; 6(1): e25-e35. <https://doi.org/10.1016/j.ecns.2009.08.004>
- [7] Lee L, Lee J, Wong K, et al. The establishment of an integrated skills training centre for undergraduate nursing education. *International Nursing Review*. 2010; 57(3): 359-364. PMID:20796066 <https://doi.org/10.1111/j.1466-7657.2009.00796.x>
- [8] Eilisha B. Effectiveness of Working Individually Versus Cooperative Groups: A Classroom-Based Research Project. University of Pennsylvania. 2007.
- [9] Adamson KA, Kardong-Edgren S, Willhaus J. An updated review of published simulation evaluation instruments. *Clinical Simulation in Nursing*. 2013; 9(9): e393-e400. <https://doi.org/10.1016/j.ecns.2012.09.004>
- [10] Zulkosky KD. Simulation use in the classroom: Impact on knowledge acquisition, satisfaction, and self-confidence. *Clinical Simulation in Nursing*. 2012; 8(1): e25-e33. <https://doi.org/10.1016/j.ecns.2010.06.003>
- [11] Jeffries PR. *The NLN Jeffries simulation theory*: Wolters Kluwer; 2016.
- [12] Tosterud R, Hall-Lord ML, Petzäll K, et al. Debriefing in simulation conducted in small and large groups-nursing students' experiences. *Journal of Nursing Education and Practice*. 2014; 4(9): 173-182. <https://doi.org/10.5430/jnep.v4n9p173>
- [13] Unver V, Basak T, Watts P, et al. The reliability and validity of three questionnaires: the student satisfaction and self-confidence in learning scale, simulation design scale, and educational practices questionnaire. *Contemporary Nurse*. 2017; 53(1): 60-74. PMID:28084900 <https://doi.org/10.1080/10376178.2017.1282319>
- [14] Jeffries P, Rogers K. Evaluating simulations. *Simulation in nursing education: From conceptualization to evaluation*. 2007; 87-103.
- [15] Jeffries PR, Rizzolo MA: Designing and implementing models for the innovative use of simulation to teach nursing care of ill adults and children: A national, multi-site, multi-method study. New York, NY: National League for Nursing 2006.
- [16] Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*.

- 2000; 25(24): 3186-3191. PMID:11124735 <https://doi.org/10.1097/00007632-200012150-00014>
- [17] Shea JA, Fortna GS. Psychometric methods. In: International handbook of research in medical education. edn.: Springer; 2002; 97-126. [https://doi.org/10.1007/978-94-010-0462-6\\_4](https://doi.org/10.1007/978-94-010-0462-6_4)
- [18] Jeffries PR, Rodgers B, Adamson K. NLN Jeffries simulation theory: Brief narrative description. *Nursing Education Perspectives*. 2015; 36(5): 292-293. PMID:26521496 <https://doi.org/10.1097/00024776-201509000-00004>
- [19] Hair JF, Anderson RE, Tatham RL, et al. Black (1998). *Multivariate data analysis*. 1998; 5: 87-135.
- [20] Nunnally JC, Bernstein IH. *Psychological theory*. In.: New York: McGraw-Hill; 1994.
- [21] Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *Journal of Evaluation in Clinical Practice*. 2011; 17(2): 268-274. PMID:20874835 <https://doi.org/10.1111/j.1365-2753.2010.01434.x>
- [22] Tabachnick BG, Fidell LS. *Principal components and factor analysis. Using Multivariate Statistics*. 2001; 4: 582-633.
- [23] Norman GR, Streiner DL. *Biostatistics: the bare essentials: PMPH USA*; 2008.
- [24] Loiacono ET, Watson RT, Goodhue DL. WebQual: A measure of website quality. *Marketing Theory and Applications*. 2002; 13(3): 432-438.
- [25] Haccoun RR. Une nouvelle technique de vérification de l'équivalence de mesures psychologiques traduites. *Revue Québécoise de Psychologie*. 1987.
- [26] Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength & Conditioning Research*. 2005; 19(1): 231-240. PMID:15705040 <https://doi.org/10.1519/00124278-200502000-00038>
- [27] Shrout PE. Measurement reliability and agreement in psychiatry. *Statistical methods in Medical Research*. 1998; 7(3): 301-317. PMID:9803527 <https://doi.org/10.1177/096228029800700306>
- [28] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*. 2016; 15(2): 155-163. PMID:27330520 <https://doi.org/10.1016/j.jcm.2016.02.012>
- [29] Polit DF, Beck CT. *Nursing research: Generating and assessing evidence for nursing practice: Lippincott Williams & Wilkins*; 2008.
- [30] Tabachnick BG, Fidell LS. *Using multivariate statistics, 6th edn* Boston, Ma: Pearson; 2013.
- [31] Field A. *Discovering statistics using SPSS:(and sex, drugs and rock'n'roll)*, vol. 497: Sage; 2000.