Scientific Research

Application of the Objective Structure Clinical Evaluation in Evaluating Clinical Competence for a BSN Program

Jun-Yu Fan, Li-Fen Chao, Sui-Whi Jane*

Department of Nursing & Graduate Institute of Nursing, Chang Gung University of Science and Technology, Kweishan, Taoyuan, Taiwan
Email: *swjane@gw.cgust.edu.tw

Abstract

Given the complex clinical situations and the dynamic nature of nursing, the greatest challenge for nursing educators is performing objective evaluation of students' clinical competence. The Objective Structured Clinical Evaluation (OSCE) is designed to address this challenge and it has been widely used in nursing education. However, its implementation in nursing education in Taiwan has been limited. Accordingly, a quasi-experimental design was utilized to examine students' clinical performance and stress levels using a 12-station OSCE assessment. Further, we investigated the inter-rater reliability and internal consistency of the OSCE. After controlling for scores of clinical performance, overall differences in pre- and post-practicum stress in the OSCE group were significantly higher than in the control group ($F(1, 89) = 4.89$, $p = 0.03$). There was no group effect on practicum performance after controlling for grade point average ($F(1, 89) = 2.69$, $p = 0.14$). Cronbach's alpha for 12 OSCE stations ranged from 0.22 to 0.80 and inter-rater reliability for all 12 stations showed Pearson's r ranging between 0.76 and 1.00. Cohen's kappa ranged from 0.70 to 1.00 ($p < 0.001$). Future studies should explore how OSCE can be best and most cost-effectively incorporated into the BSN curriculum.

Keywords

---

*Corresponding author.

1. Introduction

The nursing profession involves complex clinical situations requiring considerable adaptability, and thus most employers expect new graduates to be well prepared for a wide range of functions in addition to specific entry-level skills for providing safe care [1] [2]. To date, the literature has indicated that nursing graduates are not sufficiently competent to handle the challenges faced in daily practice [1], resulting in disappointed employers, frustrated new graduates and dissatisfied patients. Thus, it is important for nursing educators to develop pedagogy and practicum training to enhance students' clinical competence, thereby minimizing the education-service gap.

To foster professional socialization and gain confidence for nursing student, preceptorship programs in the clinical practicum have been successfully implemented to reduce anxiety and stress as well as promote smooth adaptation from student nurse to novice nurse [3]. To validate students' competence, numerous outcome evaluations have been investigated to demonstrate the abilities of new graduates, and these have been adapted by nursing education [4]. Over the past 20 years, however, the greatest challenges have been how to measure competence precisely and objectively, as the nature of competence is multi-faceted and influenced by students' level of confidence, comfort and self-efficacy [5].

The Objective Structured Clinical Examination (OSCE) has been identified as a useful assessment strategy for evaluating students' learning and clinical performance in the medical and nursing disciplines [6]-[8]. Walsh *et al.*'s [8] integrative literature review found support for the utility of OSCE as a strategy to measure clinical competence in nursing, but the greatest concerns included a paucity of studies documenting the suitability of the design for measuring nursing clinical competence and the psychometric properties of the OSCE. To date, no studies have examined the effects of utilizing the OSCE as a strategy to measure students' clinical competence prior to the implementation of preceptorship programs. Thus, the purpose of this study was to investigate the effects of OSCE before the last-mile practicum, which is the final 6-week clinical practicum that usually commences at the beginning of the second semester of a nursing student's final year when a student is assigned to a preceptor and works as a novice nurse. We also examined the reliability and validity of the OSCE at each OSCE station.

2. Background

Stress, which reflects a dynamic relationship between the individual and environment, is a multidimensional phenomenon [9]. Gibbons, Dempster and Moutray [10] suggested that an optimum amount of stress, "eustress," is necessary for mental and physical well-being. However, excessive stress can result in sleep disorders, restlessness, forgetfulness, abnormal fatigue, poor concentration, memory impairment, and hindered problem-solving [9], thus potentially impeding students' attainment of educational goals [11]. Previous literature has demonstrated that principal stressors identified by student nurses included examinations, level and intensity of academic workload, the theory–practice gap, and poor relationships with clinical staff during their study period [12]. Among these, student nurses rate preceptorship as the most stressful clinical staff relationship [13]-[15]. Additionally, the lack of experience, difficult patients, evaluation by faculty, and anxiety about dispensing the wrong medication or information foster stress among novice student nurses stressors, making practicum the most anxiety-inducing component of nursing programs [16]. Despite its considerable utility, the preceptorship is rated as the most stressful and challenging practicum experience because the preceptor and student, who are unlikely to be familiar with each other, must work together in arduous working environments [15]. The frustration associated with a difficult relationship could result in student stress, dissatisfaction in the nursing profession, and interference with learning and assimilation [15]. From a student-centric perspective, it is imperative to transform nursing education practices and reduce stressors, thereby facilitating student continuation and academic success during professional nursing development [17].

The OSCE, first described by Harden and colleagues [18], has been identified as a useful assessment strategy for evaluating students' learning and clinical performance [6] [7]. Previous research has indicated the following benefits of the OSCE in medical and nursing disciplines: 1) increased student confidence when faced with challenges in practice [19]; 2) greater objectivity in comparison to most clinical competence measures [20] [21]; 3) simultaneous testing of a large number of students on a broader range of skills and knowledge [8] [21]; 4) reduced risk of examiner bias due to a wide range of examiners [22]; 5) increased communication skills with standardized patients (SPs) [23]; 6) timely feedback regarding errors or concerns during the assessment period [24]; 7) increased skill consistency among students [8]. Despite the anxiety and stress surrounding assessment,

most students report positive perceptions of the OSCE [8]. Although studies have found anxiety to be a confounding factor for OSCE, anxiety might also positively affect performance [8] [22] and more accurately simulate real-life emergencies [6]. Walsh *et al.*'s [8] integrative literature review found support for the utility of OSCE as a strategy to measure clinical competence in medicine (*n* = 23 papers) and nursing (*n* = 18 papers). However, as mentioned in the Introduction, few studies have documented the suitability of the design for measuring nursing clinical competence and the validity and reliability of the OSCE. Given these limitations, little is currently known about the effects and quality of OSCE when implemented in the nursing discipline. Thus, the aim of this study was to examine the effects of OSCE prior to the last-mile practicum on student stress levels and clinical practicum performance. We also evaluated the quality of OSCE with the inter-rater reliability and internal consistency of each OSCE station.

3. Methods

3.1. Design, Setting, and Sample

This study employed a quasi-experimental design and was conducted in a northern Taiwanese university, from October 2010 to September 2011. The university has two campuses with approximately 7600 enrolled students in 2011 with about 600 (12 classes) BSN students. Participants were recruited using convenience sampling from final-year (*n* = 103, two classes) BSN students. Participants had to have completed their fundamental, medical-surgical, obstetric, pediatric, psychiatric, and community health nursing practicums. One student was thus excluded from the study sample, leaving a total of 102. Ten then declined participation due to a lack of interest (9.8% refusal rate), and thus, the final study sample consisted of 92 nursing students. Two weeks prior to the last-mile practicum, two class representatives selected one of two sealed envelopes in the presence of the researcher to determine the experimental (Group O; *n* = 38) and control groups (Group C; *n* = 54). Group O performed a 12-station OSCE 3 days before the last-mile practicum, while students in Group C received the standard last-mile practicum preparatory information and training.

3.2. Instruments

3.2.1. Demographic Data
Demographic data included participants' gender, age, previous 3 years' grade point average (GPA), and last-mile practicum score.

3.2.2. Perceived Stress Scale of Nursing Students in Clinical Practice
The Perceived Stress Scale of Nursing Students in Clinical Practice questionnaire was developed by Sheu *et al*. [25] to measure the types and degree of perceived stressful events, henceforth referred to as the Practicum Stress Scale (PSS). The PSS, which is widely used [13] [26] [27], contains 29 items across six dimensions of stressors: taking care of patients (eight items), school faculty and hospital staff (six items), homework and workload (five items), peers and personal life (four items), clinical environment (three items), and professional knowledge and skills (three items). Each item is scored on a 5-point Likert-type scale from 0 ("never", frequency 0% - 10%) to 4 ("all the time", frequency 90% - 100%). Total scores can range from 0 to 116 points, with higher scores indicating greater stress. Cronbach's alphas for individual subscales were between 0.59 and 0.81, and the overall Cronbach's alpha was 0.89. The 1-week test-retest reliability of 0.60 (*p* < 0.01) demonstrated the reliability of this instrument, while the content validity index of 0.94 indicated its validity. In addition, 50.7% of the total variance was accounted for by the six factors, which confirmed the construct validity of this instrument [25]. Both groups filled a practicum-related stress questionnaire 1 week before and after the last-mile practicum.

3.2.3. Practicum Performance
In general, subject grades represent the most popular indicator of academic performance in university settings. Therefore, students' practicum clinical performance was measured with a department-developed evaluation form with four domains: 1) clinical performance (50%; data collection, diagnosis, skills, and interventions); 2) communication skills (15%; patients, families, and the medical team); 3) professionalism (15%; attitude, ethics, and patient-centric care); 4) homework (20%). The clinical performance score comprised evaluation from the preceptor (weight 80%) and teacher coordinator (weight 20%).

3.3. Procedures

This study was approved by the institutional review board for studies with human subjects from Chang Gung Memorial Hospital (IRB 99-2245B). The principal investigator (PI) explained the purpose of the study to potential participants 3 months before the last-mile practicum began, after which content and procedures related to the OSCE were posted on the nursing department website. One week after PI provided this explanation, informed consent from all potential participants was obtained.

3.3.1. Preparation for the Last-Mile Practicum

The nursing school participating in this study offers a 1056-hour practicum on fundamental, medical-surgical, obstetric, pediatric, psychiatric, and community health nursing, and a "last-mile" practicum. The last-mile practicum, *i.e.*, the final 6-week clinical practicum, usually commences at the beginning of the second semester of a nursing student's final year. During the last-mile practicum, students are assigned a preceptor (senior unit staff member), preceptorship, for individual guidance. Students and incumbents have similar roles and workloads, except for the additional preceptor supervision. Additionally, a teacher coordinates for administrative purposes, communicating with the preceptor, student, and head nurse to ensure that the practicum runs smoothly. The teacher also discusses case report homework, clinic-related problems and problem-solving skills, and addresses other needs with students in a weekly, 4-hour class. Students' final clinical performance scores are given by the preceptor and teacher coordinator.

To equip students' clinical competence, the university provides nursing students with self-learning tools in fundamental and medical-surgical skills (e.g., catheterization, venipuncture, dressing changes, and intravenous injection) using high-fidelity simulators (e.g., Resusci-Anne, Laerdal™, and SimMan™) in clinical competence centers (CCC). Alternatively, students can view standard operating procedures for these skills through the multimedia on demand platform on the university's intranet system. Both groups received same preparation prior to last-mile practicum.

3.3.2. Implementation of the Objective Structured Clinical Evaluation

Brannick, Erol-Korkmaz and Prewett [28] conducted a meta-analysis describing the effects of number of stations, number of raters, competence dimensions, and measurement purpose. The overall (summary) mean alpha across stations was 0.66; the overall mean alpha within stations across items was 0.78. Further, more stations tended to produce higher reliability. Specifically, the unweighted mean alpha for OSCEs with <10 stations was 0.56, but 0.74 for OSCEs with >10 stations. In Taiwan, the OSCE was adopted by institutions in 2006, and it has been mandated as a national medical board examination since 2012. In addition, several healthcare disciplines, including Chinese medicine, dentistry, pharmacology, and nursing, have adapted it to objectively assess students' clinical competence.

The nursing discipline has not been adopted in the national board examination, and thus the OSCE designed for this study was adapted from the 12-station version (120 min) used for Taiwanese medical students. To fully reflect multidimensional clinical competences, this OSCE procedure included six key dimensions: knowledge application and skills practice (five stations), problem identification (three stations), patient and family education (one station), clinical documentation (one station), ethics and patient safety (one station), and communication skills (one station).

Two important core competences for nurses, communication skills and ethical issues, were unobtrusively incorporated into each OSCE station. The 12-station OSCE included eight stations using SPs and four using manikin scenarios. Each station took 10 min to complete: 1 min for reviewing the exam question, 1 min for moving around stations, and 8 min for completing task components. The OSCE was conducted 3 days prior to the last-mile practicum at the clinical competence center (CCC) in the nursing department, a setting accredited by the Taiwanese Association of Medical Education and qualified for the national medical board examination.

Ten non-nursing students and school staff from the university trained by experienced faculty members were recruited as SPs. At each examination station, a checklist was developed to assess desired behaviors or competences, with 18 - 30 checklist items. Another two expert faculty members reviewed the checklist for accuracy and thoroughness. Prior to the formal OSCE, two junior students completed a 12-station OSCE to ensure the clarity of questions and sufficiency of allotted time at each station. Further, all 24 raters (two raters for each station) attended the pilot run to familiarize themselves with all OSCE procedures and reach consensus on checklist

scoring.

The total duration of OSCE was 120 min per student, and it was conducted in morning and afternoon sessions. Prior to the OSCE, a co-author explained the OSCE process and a staff member guided the first 12 students to their designated stations. After completing the OSCE, the students completed the PSS and a co-author shared their most common mistakes and major errors during the OSCE. To reduce contamination, students were not allowed to leave until all 38 had completed the OSCE.

Students in Group C received the standard last-mile practicum preparatory information and training, but not the 12-station OSCE. Therefore, they completed the PSS 1 week prior to the last-mile practicum. Both groups completed the PSS 1 week after the last-mile practicum.

3.4. Data Analysis

Data were analyzed using SPSS 17.0 for Windows (SPSS Inc., Chicago, IL, USA), and the threshold for statistical significance was set at $p < 0.05$. Prior to data analysis, we confirmed that data were appropriate for the chosen statistical analyses. Descriptive statistics, independent and paired *t*-tests, analyses of variance (ANOVA), and analyses of covariance (ANCOVA) were conducted to test for differences between groups and a potential main (group) effect on practicum performance after controlling for covariates. Cronbach's alpha, Pearson's *r*, and Cohen's kappa were used to examine the internal consistency of each station and inter-rater reliability.

4. Results

4.1. Demographics

All 92 participants were female, aged between 21.13 - 25.58 years ($M = 22.00 \pm 0.56$). More than half ($n = 56$, 60.90%) were completing their practicum in a medical center with more than 5000 beds. The general ward was the most common practicum unit ($n = 42$, 45.70%), followed by emergency or intensive care ($n = 39$, 42.4%) and operating room ($n = 11$, 12.00%) (**Table 1**). In testing for homogeneity between groups, GPA for the previous 3 years was found to be significantly lower in Group O than in Group C ($t = -2.09$, $p = 0.039$) after applying Levene's test for equality of variances. There were no significant differences between groups on the six subscales of practicum stress before the practicum.

4.2. Practicum Stress

A univariate ANCOVA was used to examine whether there was a main group effect on level of posttest practicum stress after the last-mile practicum after controlling for three potential confounding variables—level of pretest practicum stress, clinical performance, and GPA for the previous 3 years—to adjust for baseline and homogeneity differences between groups (**Table 1**). There was a significant group effect on last-mile practicum stress, ($F (1, 87) = 8.86$, $p = 0.004$) with lower stress in Group O (1.36 ± 0.57) than in Group C (1.59 ± 0.54). Pretest practicum stress ($F (1, 87) = 8.39$, $p = 0.005$) and practicum performance ($F (1, 87) = 20.00$, $p < 0.001$) significantly affected posttest practicum stress. Further analysis of six stress-inducing factors by controlling for the above three covariates indicated that the decrease in stress from "taking care of patients", "homework and workload", and "professional knowledge and skills" for Group O was significantly greater than it was for Group C, $F (1, 87) = 8.38$, $p = 0.005$; $F (1, 87) = 16.61$, $p < 0.001$; and $F (1, 87) = 7.30$, $p = 0.008$, respectively (**Table 2**).

4.3. Practicum Performance

We employed a univariate ANCOVA to examine whether there was a group effect on practicum performance after controlling for GPA for the previous 3 years and level of pretest practicum stress. Results indicated there was no group effect ($F (1, 87) = 2.63$, $p = 0.108$) on practicum performance, but GPA for the previous 3 years was found to significantly affect practicum performance ($F (1, 87) = 5.74$, $p = 0.019$).

4.4. Internal Consistency of OSCE Stations and Inter-Rater Reliability

Each OSCE station can be considered analogous to a scale or domain of an inventory. Thus, Cronbach's alpha, a measure of internal consistency reliability, was appropriate for examining whether students in identical scenarios consistently produced similar results. Thus, Cronbach's alpha for each station was calculated to ensure in-

Table 1. Descriptive statistics.

Characteristics	M ± SD or No. (%)		
	OSCE (*n* = 38)	Control (*n* = 54)	Total (*n* = 92)
Age (years)	21.96 ± 0.37	22.03 ± 0.66	22.00 ± 0.56
Type of hospital			
Medical center	24 (63.20)	32 (59.30)	56 (60.90)
Regional hospital	14 (36.80)	22 (40.70)	36 (39.10)
Practicum unit			
General ward	13 (34.20)	29 (53.70)	42 (45.70)
ED/ICU	20 (52.60)	19 (35.20)	39 (42.40)
OR	5 (13.20)	6 (11.10)	11 (12.00)
GPA for previous 3 years*	80.97 ± 2.90	82.56 ± 4.38	81.90 ± 3.90
Total pre-test practicum stress score	60.63 ± 20.73	59.78 ± 15.24	60.13 ± 17.61
Mean pre-test practicum stress score	2.09 ± 0.71	2.06 ± 0.53	2.07 ± 0.61
Taking care of patients	2.26 ± 0.74	2.10 ± 0.62	2.16 ± 0.67
School faculty and hospital staff	1.72 ± 0.86	1.83 ± 0.60	1.79 ± 0.72
Homework and workload	2.48 ± 0.78	2.43 ± 0.66	2.46 ± 0.71
Peers and personal life	1.76 ± 0.87	1.73 ± 0.71	1.74 ± 0.78
Professional knowledge and skills	2.20 ± 0.73	2.28 ± 0.72	2.25 ± 0.72
Clinical environment	2.04 ± 0.92	2.01 ± 0.75	2.03 ± 0.82
Mean posttest practicum stress score	1.36 ± 0.57	1.59 ± 0.54	1.50 ± 0.56
Practicum performance score (Clinical portion)	66.96 ± 3.50	67.59 ± 2.93	67.33 ± 3.18
Total practicum performance score*	84.00 ± 4.25	85.67 ± 3.30	84.98 ± 3.79

Note: OSCE: Objective Structured Clinical Examination; SD: Standard deviation; NTD: New Taiwan Dollar; ED: Emergency department; ICU: Intensive care unit; OR: Operating room; *$p < 0.05$.

Table 2. Comparison of mean stress scores before and after last-mile practicum.

	Group O (*n* = 38)		Group C (*n* = 54)		Differences Group O/C	
	Mean ± SD	*t* paired	Mean ± SD	*t* paired	Mean ± SD	ANCOVA
Mean practicum stress scores before/after	2.09 ± 0.71/1.36 ± 0.57	6.19***	2.06 ± 0.53/1.59 ± 0.54	6.52***	0.73 ± 0.73/0.47 ± 0.53	4.89*
Taking care of patients	2.26 ± 0.74/1.54 ± 0.64	6.30***	2.10 ± 0.62/1.73 ± 0.54	4.99***	0.72 ± 0.70/0.37 ± 0.54	8.24**
School faculty and hospital staff	1.72 ± 0.86/1.22 ± 0.71	2.97**	1.83 ± 0.60/1.33 ± 0.65	4.80***	0.50 ± 1.05/0.50 ± 0.76	0.184
Homework and workload	2.48 ± 0.78/1.41 ± 0.74	7.10***	2.43 ± 0.66/1.89 ± 0.78	4.94***	1.08 ± 0.94/0.54 ± 0.80	10.38**
Peers and personal life	1.76 ± 0.87/1.11 ± 0.73	4.23***	1.73 ± 0.71/1.26 ± 0.79	5.03***	0.65 ± 0.95/0.47 ± 0.69	0.13
Professional knowledge and skills	2.20 ± 0.73/1.54 ± 0.74	5.00***	2.28 ± 0.72/1.91 ± 0.66	3.84***	0.66 ± 0.81/0.37 ± 0.71	3.05
Clinical environment	2.04 ± 0.92/1.25 ± 0.68	4.62***	2.01 ± 0.75/1.32 ± 0.82	6.15***	0.79 ± 1.05/0.69 ± 0.83	0.41

Notes: ANCOVA: analysis of covariance; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

ternal consistency within each checklist. Acceptable internal consistency is indicated by a Cronbach's alpha value between 0.60 and 0.70, good internal consistency by values between 0.70 and 0.90 and excellent by values above 0.90 [29]. Internal consistency for stations 1, 3, 4, and 7 - 12 was acceptable to good (0.63 - 0.80), but poor for stations 5 and 6 (0.53 - 0.59), and unacceptable for station 2 (0.22).

Regarding inter-rater reliability of each station across all 12 stations, an independent samples *t*-test was conducted to examine whether there were differences between the two raters at each station. There were no inter-rater differences across the 12 stations (*p*s > 0.005). Prior to exploring the relationship between any two raters

for each participant's exam score at each station, the Shapiro-Wilk test was used to determine the normality of the mean difference scores at each station. The normality assumption was found to be violated at stations 6, 7, 9, and 11, necessitating the use of a non-parametric test, Spearman's rho, to test the relationship between the two raters at stations 6, 7, 9, and 11 with Spearman's $\rho s = 0.97$. Pearson's *r* was used to explore the relationship between any two raters for each participant's exam score for the remaining stations. Significant correlations were observed across other 8 stations, (rs = 0.76 - 1.00). We created a four-level overall performance scoring system with a straightforward 0 - 100 numeric scoring system, regrouping scores into "excellent" (≥ 90), "good" (≥ 80 and <90), "average" (≥ 60 and <80), and "failing" (<60). Cohen's kappa used to examine inter-rater reliability between two raters for categorized data, and these values ranged from 0.70 to 1.00 (κs < 0.001), except for station 10, indicating substantial agreement across almost all stations (**Table 3**).

5. Discussion

5.1. Practicum Stress

The reduction in last-mile practicum stress levels was significantly greater for Group O (from 2.09 ± 0.71 to 1.36 ± 0.57) than for Group C (from 2.06 ± 0.53 to 1.59 ± 0.54). However, of the three covariates, level of pretest practicum stress and practicum performance significantly affected posttest practicum stress. Decreases in three stress-causing factors, namely, "taking care of patients", "homework and workload", and "professional knowledge and skills" were significantly greater in Group O. In the present study, "homework and workload" was rated as the most stressful dimension prior to the clinical practicum over "relationships with clinical staff," a finding that differed from most previous studies [14] [15]. However, it was consistent with Evans and Kelly [12], who indicated level and intensity of academic workload as primary stressors. Considering that the last-mile practicum (one preceptor for each student) differs completely from students' previous clinical experiences (one clinical instructor for every seven students) offering new sources of stress, it is understandable that students from both groups might find working independently to be difficult. "Taking care of patients" and "professional knowledge and skills" were among the top three stressors before and after the practicum for both groups, consistent with three previous studies on sources of stress associated with clinical practice [26] [30] [31]. However, Group O had lower levels of stress associated with the three stress-causing factors than did Group C. This may have been facilitated by the increased confidence, communication skills, and techniques fostered by the preparatory OSCE evaluation [19] [23] [24].

5.2. Practicum Performance

There were no significant differences between the two groups in practicum performance grades after controlling for GPA and pre-practicum stress level. This might be because 1) practicum performance grades were scored by different preceptors with a wide degree of variation; 2) tasks completed during the practicum differed considerably, with different areas (e.g., operating room, intensive care unit, emergency department, general ward, rehabilitation unit, and nursing home) entailing very different tasks, workload demands, and work environments; and/or 3) the student-preceptor relationship is an important determinant of stress during one's practicum [30]-[32], and frustrating or difficult relationships might result in student stress, dissatisfaction in the nursing profession, and/or interference with learning and clinical performance [15]. The present study did not establish predictive validity; students who performed the OSCE obtained identical practicum performance grades as the control group. Thus, this study indicates that overall OSCE performance may not fully reflect clinical competence as measured by practicum performance, and it may not predict the transferability of nursing knowledge and skill learned from in-class settings to real clinical situations [28]. Future studies might consider employing measurements that specifically examine clinical competence with direct observation, including assessments and a portfolio using a valid nurse competency inventory [1] [5].

5.3. Quality of the OSCE

The overall mean alpha across stations was 0.64 (0.22 - 0.80), which was lower than 0.78 reported by Brannick *et al.* [28]. The unstable internal consistency of OSCE stations might be due an excess of checklist items (18 - 30 per station) and complex scoring systems, particularly at the skill practice station.

According to most literature indicating good agreement with inter-rater scores $\gamma = 0.6$ [20] reliability of 0.76 -

Table 3. Content validity and inter-rater reliability among the 12 stations.

Dimension	Station	Content Validity	Mean Scores of Raters 1 and 2		Inter-Rater Reliability		
		Cronbach's Alpha	Rater 1 M ± SD	Rater 2 M ± SD	*p*-Value	Pearson's *r*/ Spearman's ρ	Kappa κ
Knowledge application & skills practice	1. Vital signs	0.70	59.68 ± 19.94	59.89 ± 20.01	0.963	1.00**	1.00***
Knowledge application & skills practice	2. Medication: oral administration	0.22	39.53 ± 14.77	38.68 ± 16.27	0.814	0.87**	1.00***
Knowledge application & skills practice	3. Intravenous drip	0.80	40.37 ± 17.11	40.53 ± 17.39	0.968	1.00**	1.00***
Knowledge application & skills practice	4. Suction	0.73	61.97 ± 15.88	63.11 ± 15.88	0.757	0.97**	0.92***
Communication skills	5. History taking	0.59	49.68 ± 12.91	49.82 ± 13.73	0.966	0.97**	0.93***
Patient and family education	6. Fever routine	0.53	30.82 ± 9.68	31.39 ± 10.53	0.799	0.97**a	1.00***
Problem identification	7. Abdomen assessment	0.73	27.24 ± 19.34	26.45 ± 19.38	0.859	0.97**a	1.00***
Problem identification	8. Pain assessment	0.69	53.09 ± 15.67	52.47 ± 15.71	0.864	0.92**	0.70***
Knowledge application & skills practice	9. Catheterization	0.72	66.11 ± 14.76	65.24 ± 15.77	0.805	0.97**a	0.88***
Clinical documentation	10. Change dressing	0.63	67.82 ± 9.72	67.26 ± 8.12	0.781	0.76**	0.24*
Problem identification	11. Finger prick & hyperglycemia	0.65	70.26 ± 12.26	71.05 ± 11.99	0.777	0.97**a	0.91***
Ethics/patient safety	12. Operational permission	0.70	47.53 ± 23.55	48.53 ± 21.76	0.848	0.98**	1.00***

aSpearman's rho; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

1.00 for all 12 stations in this study was regarded as good reliability of OSCE and that was consistent with one study [33] (history-taking 0.73 - 0.96 and neurological physical examination abilities 0.84 - 0.8, and better than another (0.53 - 0.96) [34].

In conclusion, implementation of the OSCE prior to the beginning of the last-mile practicum in student nurses had significant effects on reducing practicum stress. However, the effects of the OSCE on clinical competence were limited. From measures of both educators and students, however, the OSCE appears to be a valid, reliable, and comprehensive instrument for assessing nursing skills in conjunction with clinical practice and evaluation. Still, OSCE evaluation cannot replace experience gained from clinical settings [7]. Genuine clinical experiences have always been a key component of healthcare education, and the OSCE can supplement evaluation in nursing education. Future studies might consider exploring cost-effective and efficient methods of incorporating the OSCE into the nursing curriculum for the improvement of evaluation.

5.4. Limitations and Suggestions

This study had a relatively small sample of nursing students and was geographically restricted to one teaching university in northern Taiwan, perhaps limiting its generalizability. Future randomized clinical trials are recommended to decrease sampling bias. Given that competence will remain key in nursing practice and education for the foreseeable future, multidimensional measurement of clinical competence is recommended. Due to the accumulative nature of clinical competence, future studies may implement a longitudinal design to lengthen the observation time period and implement the OSCE at the beginning of the BSN program rather than in the final year.

In terms of the quality of the OSCE in this study, an excess of checklist items per station, different scoring criteria among stations and the resulting wide range of scores may have reduced indices of content validity. In future studies, we suggest establishing more widely applied standard OSCE protocols, such as <15 checklist items per station and rules for awarding and deducting points for answers. Furthermore, the cost of implementing the OSCE is considerable. Cost can vary depending on the number of stations, raters per station and SP, yet OSCE remains a costly evaluation tool compared with traditional paper and pencil tests. Despite these limita-

tions, cost-effective incorporation of the OSCE into the nursing curriculum could help build clinical competence and create links between learning and performance in student nursing assessments.

Acknowledgements

References

[1] Garside, J.R. and Nhemachena, J.Z.Z. (2013) A Concept Analysis of Competence and Its Transition in Nursing. *Nurse Education Today*, **33**, 541-545. http://dx.doi.org/10.1016/j.nedt.2011.12.007

[2] Tilley, D.S. (2008) Competency in Nursing: A Concept Analysis. *Journal of Continuing Education in Nursing*, **39**, 58-64. http://dx.doi.org/10.3928/00220124-20080201-12

[3] Bourbonnais, F.F. and Kerr, E.J. (2007) Preceptoring a Student in the Final Clinical Placement: Reflections from Nurses in a Canadian Hospital. *Journal of Clinical Nursing*, **16**, 1543-1549. http://dx.doi.org/10.1111/j.1365-2702.2006.01828.x

[4] Anema, M. and McCoy, J. (2010) Competency-Based Nursing Education: Guide to Achieving Outstanding Learner Outcomes. Springer, New York.

[5] Pijl-Zieber, E.M., Barton, S., Konkin, J. Awosoga, O. and Caineet, V. (2014) Competence and Competency-Based Nursing Education: Finding Our Way through the Issues. *Nurse Education Today*, **34**, 676-678. http://dx.doi.org/10.1016/j.nedt.2013.09.007

[6] Byrne, E. and Smyth, S. (2008) Lecturers' Experiences and Perspectives of Using an Objective Structured Clinical Examination. *Nurse Education Practice*, **8**, 283-289. http://dx.doi.org/10.1016/j.nepr.2007.10.001

[7] Houghton, C.E., Casey, D., Shaw, D. and Murphy, K. (2012) Staff and Students' Perceptions and Experiences of Teaching and Assessment in Clinical Skills Laboratories: Interview Findings from a Multiple Case Study. *Nurse Education Today*, **32**, e29-e34. http://dx.doi.org/10.1016/j.nedt.2011.10.005

[8] Walsh, M., Bailey, P.H. and Koren, I. (2009) Objective Structured Clinical Evaluation of Clinical Competence and Integrative Review. *Journal of Advanced Nursing*, **65**, 1584-1595. http://dx.doi.org/10.1111/j.1365-2648.2009.05054.x

[9] Beddoe, A.E. and Murphy, S.O. (2004) Does Mindfulness Decrease Stress and Foster Empathy among Nursing Students? *Journal of Nursing Education*, **43**, 305-312.

[10] Gibbons, C., Dempster, M. and Moutray, M. (2007) Stress and Eustress in Nursing Students. *Journal of Advance Nursing*, **61**, 282-290. http://dx.doi.org/10.1111/j.1365-2648.2007.04497.x

[11] Del Prato, D., Bankert, E., Grust, P. and Joseph, J. (2011) Transforming Nursing Education: A Review of Stressors and Strategies That Support Students' Professional Socialization. *Advances in Medical Education and Practice*, **6**, 109-116. http://dx.doi.org/10.2147/AMEP.S18359

[12] Evans, W. and Kelly, B. (2004) Pre-Registration Diploma Student Nurse Stress and Coping Measures. *Nurse Education Today*, **24**, 473-482. http://dx.doi.org/10.1016/j.nedt.2004.05.004

[13] Chen, Y.-W. and Hung, C.-H. (2014) Predictors of Taiwanese Baccalaureate Nursing Students' Physio-Psycho-Social Responses during Clinical Practicum. *Nurse Education Today*, **34**, 73-77. http://dx.doi.org/10.1016/j.nedt.2013.02.021

[14] Duchscher, J.E.B. (2009) Transition Shock: The Initial Stage of Role Adaptation for Newly Graduated Registered Nurses. *Journal of Advance Nursing*, **65**, 1103-1113. http://dx.doi.org/10.1111/j.1365-2648.2008.04898.x

[15] Yonge, O., Myrick, F. and Haase, M. (2002) Student Nurse Stress in the Preceptorship Experience. *Nurse Educator*, **27**, 84-88. http://dx.doi.org/10.1097/00006223-200203000-00012

[16] Sharif, F. and Masoumi, S. (2005) A Qualitative Study of Nursing Student Experiences of Clinical Practice. *BMC Nursing*, **9**, 4-6.

[17] Hopkins, T.H. (2008) Early Identification of At-Risk Nursing Students: A Student Support Model. *Journal of Nursing Education*, **47**, 254-259. http://dx.doi.org/10.3928/01484834-20080601-05

[18] Harden, T.K., Klein, R.L., Smith, R.R., Thureson-Klein, A.K. and Lowry, M.W. (1975) Serum Dopamine Beta-Hydroxylase Activity Following Subarachnoid Hemorrhage in Man. *Medical Biology*, **53**, 100-106.

[19] Alinier, G. (2003) Nursing Students' and Lecturers' Perspectives of Objective Structured Clinical Examination Incorporating Simulation. *Nurse Education Today*, **23**, 419-426. http://dx.doi.org/10.1016/S0260-6917(03)00044-3

[20] Schuwirth, L.W. and van der Vleuten, C.P. (2003) The Use of Clinical Simulations in Assessment. *Medicine Education*, **37**, 65-71. http://dx.doi.org/10.1046/j.1365-2923.37.s1.8.x

[21] Watson, R., Stimpson, A., Topping, A. and Porock, D. (2002) Clinical Competence Assessment in Nursing: A Systematic Review of the Literature. *Journal of Advanced Nursing*, **39**, 421-431. http://dx.doi.org/10.1046/j.1365-2648.2002.02307.x

[22] Bartfay, W.J., Rombough, R., Howse, E. and Leblanc, R. (2004) Evaluation: The OSCE Approach in Nursing Education. *The Canadian Nurse*, **100**, 18-23.

[23] Hale, L.S., Lewis, D.K., Eckert, R.M., Wilson, C.M. and Smith, B.S. (2006) Standardized Patients and Multidisciplinary Classroom Instruction for Physical Therapist Students to Improve Interviewing Skills and Attitudes about Diabetes. *Journal of Physical Therapy Education*, **20**, 22-27.

[24] Nestela, D., Kneeboneb, R., Nolanb, C., Akhtarc, K. and Darzib, A. (2011) Formative Assessment of Procedural Skills: Students' Responses to the Objective Structured Clinical Examination and the Integrated Performance Procedural Instrument. *Assessment & Evaluation in Higher Education*, **36**, 171-183. http://dx.doi.org/10.1080/02602930903221469

[25] Sheu, S., Lin, H.S., Hwang, S.L., Yu, P.J., Hu, W.Y. and Lou, M.F. (1997) The Development and Testing of a Perceived Stress Scale for Nursing Students in Clinical Practice. *The Journal of Nursing Research*, **5**, 341-352.

[26] Sheu, S., Lin, H.S. and Hwang, S.L. (2002) Perceived Stress and Physio-Psycho-Social Status of Nursing Students during Their Initial Period of Clinical Practice: The Effect of Coping Behaviors. *International Journal of Nursing Studies*, **39**, 165-175. http://dx.doi.org/10.1016/S0020-7489(01)00016-5

[27] Al-Zayyat, A.S. and Al-Gamal, E. (2014) Perceived Stress and Coping Strategies among Jordanian Nursing Students during Clinical Practice in Psychiatric/Mental Health Courses. *International Journal of Mental Health Nursing*, **23**, 326-335. http://dx.doi.org/10.1111/inm.12054

[28] Brannick, M.T., Erol-Korkmaz, H.T. and Prewett, M. (2011) A Systematic Review of the Reliability of Objective Structured Clinical Examination Scores. *Medical Education*, **45**, 1181-1189. http://dx.doi.org/10.1111/j.1365-2923.2011.04075.x

[29] Kline, P. (2000) The Handbook of Psychological Testing. 2nd Edition, Routledge, London.

[30] Chan, C.K., So, W.K. and Fong, D.Y. (2009) Hong Kong Baccalaureate Nursing Students' Stress and Their Coping Strategies in Clinical Practice. *Journal of Professional Nursing*, **25**, 307-313. http://dx.doi.org/10.1016/j.profnurs.2009.01.018

[31] Jimenez, C., Navia-Osorio, P.M. and Diaz, C.V. (2010) Stress and Health in Novice and Experienced Nursing Students. *Journal of Advance Nursing*, **66**, 442-455. http://dx.doi.org/10.1111/j.1365-2648.2009.05183.x

[32] Gorostidi, X.Z., Egilegor, X.H., Erice, M.J.A., Iturriotz, M.J.U., Garate, I.E., Lasa, M.B. and Cascante, X.S. (2007) Stress Sources in Nursing Practice. Evolution during Nursing Training. *Nurse Education Today*, **27**, 777-787. http://dx.doi.org/10.1016/j.nedt.2006.10.017

[33] Jain, S.S., DeLisa, J.A., Eyles, M.Y., Nadler, S., Kirshblum, S. and Smith, A. (1998) Further Experience in Development of an Objective Structured Clinical Examination for Physical Medicine and Rehabilitation Residents. *American Journal of Physical Medicine & Rehabilitation*, **77**, 306-310. http://dx.doi.org/10.1097/00002060-199807000-00009

[34] Roberts, J. and Brown, B. (1990) Testing the OSCE: A Reliable Measurement of Clinical Nursing Skills. *The Canadian Journal of Nursing Research*, **22**, 51-59.

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or Online Submission Portal.