Implementation of Objective Structured Clinical Examination for Assessing Nursing Students’ Clinical Competencies: Lessons and Implications*

Patricia Katowa-Mukwato, Lonia Mwape, Marjorie Kabinga-Makukula, Prudencia Mweemba, Margaret C. Maimbolwa
Department of Nursing Sciences, School of Medicine, University of Zambia, Lusaka, Zambia
Email: patriciakatowamukwato@gmail.com

Received August 8th, 2013; revised September 8th, 2013; accepted September 15th, 2013

Copyright © 2013 Patricia Katowa-Mukwato et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Objective Structured Clinical Examination (OSCE) as a performance-based assessment method is a well established student assessment tool. Its popularity in the assessment of clinical competence is well documented and prominent in situations where reliability and content validity are fundamental. In this paper, we describe the implementation of OSCE in the Department of Nursing Sciences; University of Zambia for assessment of nursing students’ clinical competencies. The implementation process followed an eight step-approach from which several lessons were drawn and implications were generated. Major lessons included the need for adequate preparation of faculty and students, which is a fundamental ingredient to ensure reliability of the examination, and in minimizing stress and anxiety respectively. Following the implementation we acknowledged that OSCEs are suitable for testing clinical, technical and practical skills which may not be adequately assessed through traditional assessment methods as they possess the ability to improve the validity and reliability of assessments. Nevertheless, careful consideration should be taken to avoid entirely relying on OSCE as the only means of assessing clinical competencies.

Keywords: Clinical Competence; Objective Structured Clinical Examination; Assessment; Nursing Students

Introduction

Clinical Competence is a complex concept and debates continue about the most appropriate definition and method of assessment (Evans, 2008). Watson et al., (2002) suggest that competence is a nebulous concept defined in deferent ways by different people. Its relationship with other concepts such as capability, performance, proficiency and expertise make it even more difficult to define. Earlier Gonzi (1994) described three ways of understanding competence: 1) tasks related skills, 2) patterning to generic attributes essential to performance and 3) the bringing together of a range of general attributes such as knowledge, skills and attitudes appropriate for professional practice. Later the Australian Nurses and Midwifery Council (2005) described competence in a more holistic way as a combination of skills, knowledge, attitudes, values and abilities that underpin effective and/or superior performance in a profession.

The above definitions underscore the complexity and multifacetedness of clinical competence. The complex nature of clinical competence consequently poses a challenge in isolating or identifying suitable assessment methods that are able to measure all its attributes as well as maintaining validity, reliability and objectivity. Affirming the challenges in assessing clinical competence for nursing students, Levette-Jones and others (2010) stated that the challenge of validity, reliability, subjectivity and bias in measuring clinical competence has confronted universities for many years.

Since inception in 1978, the Department of Nursing Sciences at the University of Zambia, utilized Direct Observation of Procedural Skills (DOPS) for assessment of nursing students’ clinical competences for both formative and summative purposes. DOPS is a method for assessing procedural competence through direct observation by faculty (Holmboe & Hawkins, 2008). It was considered sufficient in assessment of clinical competence as the Department of sorely admitted Registered Nurses with diploma to upgrade to Bachelor’s Degree. These students were already practicing nurses and had been certified competent to practice nursing by the regulatory body (General Nursing Council of Zambia). Some students had also attained post registration qualifications such as Midwifery, Operating Theatre Nursing and Mental Health Nursing. Using DOPs, each student was assessed on one procedure deemed appropriate by the examining faculty. Selection of procedures was solely determined by the examiner as well as the availability of patients requiring such a procedure, as opposed to curricular core competencies and examination blue prints. There was often a lack of transparency about the objectives of the assessment and the competencies required to succeed (Marwaha, 2011). In addition, the lack of a clear marking system resulted in variability between examiners. Although DOPS was considered feasible and
acceptable by faculty and students, its inherent characteristics made it fail to meet the principles of reliability, content validity and standardization.

In 2010 the Department of Nursing Sciences implemented a competence-based curriculum and admitted the first cohort of pre-service students. Consequently it became necessary to review the clinical assessment methods to facilitate the implementation of those techniques that are authentic for measuring and enhancing clinical competence. As a result, in the 2012/2013 academic year, the department implemented the Objective Structured Clinical Examination (OSCE) for assessment of clinical competence. OSCE is a comprehensive, systematic and objective method of evaluation that involves an individual student rotating through a number of practical and theoretical “stations” where they are assessed using a set criteria (Bhat & Anald, 2006; Gormley, 2011). Byrne and Smyth (2008) described OSCE as an approach to students’ assessment in which aspects of clinical competence are evaluated in a comprehensive, consistent and structured manner, with close attention to the objectivity of the process.

OSCE was originally conceptualized by Harden and Gleeson at the University of Dundee in 1975 as a solution to the difficulties of adequate sampling and standardization common with traditional assessment methods (Wass et al., 2001; Turner & Dankoski, 2008; Marwaha, 2011). In the seminal paper on OSCEs, Harden et al., (1975) recommended OSCE as an alternative to traditional clinical testing methods due to its objectivity, reliability and ability to test multiple competencies in a controlled standardized manner, thus eliminating non-candidate variance in results (Marwaha, 2011). Although OSCE evolved from medical education, it has been used extensively in nursing worldwide (Shadia et al., 2010). It allows for testing of a wide range of knowledge, skills and attitudes and can accommodate large numbers of examinees in one examination session (Bhat & Anald, 2006; Shadia et al., 2010). It is also accepted as a fit-for-purpose instrument for measuring clinical reasoning skills (Ahmed, 2009). When compared to traditional methods of clinical skills assessment, OSCE has several advantages, and it meets the two cardinal criteria of an effective assessment tool vis-à-vis validity and reliability (Gormely, 2011; Turner & Dankoski, 2008; Auewarakul et al, 2005).

Validity refers to the extent to which an instrument measures the construct of interest (Evans, 2008), while reliability is the consistence of examinee scores over time (Turner & Dankoski, 2008). Although some studies have reported low reliability for OSCE, various methods have been found to increase its reliability; large number of stations (at least 10), large number of raters, good standardization of patients and adequate test length (at least 3 - 4 hours) to obtain a reliability of .85 - .90 (Turner & Dankoski, 2008). Bhat and Anald (2006) assert that the ability to test a wide range of knowledge, skills and attitudes in a single OSCE helps to ensure content validity which according to Downing (2003) is the most essential of the three types of validity: content, criterion and construct validity.

Despite its positive attributes the cost of implementing OSCE is high in terms of personnel, facilities, finances and time for examinees, Standardized Patients and faculty (Evans, 2008; Turner & Dankoski, 2008; Marwaha, 2011). Some studies have reported OSCE settings to be stressful or intimidating for participants although none has compared the level of stress to other forms of formal examinations (Evans, 2008). Notwithstanding the cost and time constraints, OSCEs have the capacity to improve the validity and reliability of assessments of many aspects of clinical competence. There is also literature to support implementation of OSCE as its running costs are outweighed by the benefits. For example, it has been affirmed that the running cost of OSCE is outweighed by the educational benefits as well as student satisfaction (Selim et al., 2011). In addition, OSCE has been supported as an appropriate method in evaluating nursing clinical skills because of various advantages such as, improving student clinical performance, preparing highly qualified and competent graduates, increasing decision making abilities, and enhancing teaching level (El Darir & Abd El Hamid, 2013). Furthermore, OSCEs have been reported to be beneficial as they enhance skill acquisition through hands-on approach and affords students to practice in a safe controlled environment (Evans, 2008).

**Implementation Process**

The implementation of OSCE was preceded by a series of in-house planning and orientation meetings for faculty and Staff Development Fellows (SDF) in the Department of Nursing Sciences at the University of Zambia, between January, and April 2013. Since OSCE was a new concept for most members of staff and SDFs, initial meetings were designed to give an overview of OSCE as an assessment tool, its characteristics in comparison with other clinical assessment tools and in meeting the criteria of validity and reliability. Other meetings dealt with the design of OSCE; including types and numbers of stations. Meetings were also held to discuss and agree on the steps towards developing an OSCE.

In literature, a variety of models have been used in the implementation of OSCE. For example El Darir and Abd el Hamid (2013) utilized a three step model which included construction of OSCE schemes and clinical scenario, as first phase, actual conducting of OSCE as second phase, and Evaluation as third and last phase. In our case eight steps were followed during the planning and implementation phases as outlined in Table 1.

Following consensus on the outlined steps, four teams were constituted based on subject matter expertise and included; clinical nursing, mental health and psychiatric nursing, maternal and child health nursing and community health nursing.

**Table 1.**

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>Identification of competencies to be assessed (drawn from the curriculum)</td>
</tr>
<tr>
<td>Step 2</td>
<td>Development of case scenarios based on identified competences</td>
</tr>
<tr>
<td>Step 3</td>
<td>Identification/modification/development of evaluation tools (checklist, rating scales etc.)</td>
</tr>
<tr>
<td>Step 4</td>
<td>Identification of assessment sites class rooms and clinical skills laboratory</td>
</tr>
<tr>
<td>Step 5</td>
<td>Planning for resources (human, simulator/models, medical surgical supplies and stationery)</td>
</tr>
<tr>
<td>Step 6</td>
<td>Orientation of standardized patients SPs (graduate students and support staff)</td>
</tr>
<tr>
<td>Step 7</td>
<td>Mock OSCE</td>
</tr>
<tr>
<td>Step 8</td>
<td>Implementation (actual conduct of OSCE)</td>
</tr>
</tbody>
</table>
teams developed scenarios, checklists and identified required resources, both human and material, after which all developed scenarios and checklist were reviewed by a combined team drawn from the four subject areas. Given that this was the first OSCE, students were given an orientation on the structure and what they should expect and what was expected of them.

To test the practicality of OSCE, a mock was conducted three days prior to the date of examination. The mock OSCE was used to pilot some of the scenarios and checklists and determine whether the time allocated was adequate for the performance of different skills. The mock OSCE was also used to determine inter-rater variability and provided a platform for consensus building regarding the scoring system and allocation of marks. During the mock OSCE, SDFs acted as students while faculty acted as examiners and standardized patients. Two days before the examination, some postgraduate students were oriented to act as standardized patient during the main OSCE. Finally, the first ever OSCE in the department was administered on 104 fourth year Nursing students during the 2012/2013 end of academic year examination. OSCE was used to assess clinical competences in six courses; Medical Surgical Nursing, Community Health Nursing, Maternal and Child Health Nursing, Psychiatric and Mental Health Nursing, Peaditics and Operating Theatre Nursing.

Many variants exist in the implementation of OSCE. For example, stations may be much longer and examiners may not be present, with the marking being undertaken by the simulated patients on whom the task was performed. In other situations, there may be stations at which multiple-choice questions are asked or at which other forms of written responses are required, while other stations require performance of a clinical procedure while being observed and evaluated by faculty using a standard checklist (Newbel, 2004; El Darir & Abd El Hamid, 2013). In the case of the department of Nursing Sciences, seven stations were designed for each of the six courses. Four of the seven were marker station while three were observer station. Observer stations consisted of a task presented in two to three sentence scenarios and a request for appropriate action or performance rated by an examiner using predetermined checklist. Observer stations were used mainly for evaluating skills of psychomotor nature eg measuring vital signs, physical examination, and intramuscular drug administration. Marker stations consisted of presentation of data with a request for interpretation, documentation or appropriate clinical action (Amin & Hoon-Eng, 2003). Each station was allocated 10 minutes.

As opposed to DOPS, where each student is examined on only one randomly selected skill by one examiner, with the implementation of OSCE, for the first time in the Department of Nursing Sciences, all students were examined on seven different core competencies as specified in the curriculum. In addition each student was scored by three examiners on the three observer stations. This to a large extent eliminated personal biases that usually arise from single examiners. In addition, the use of predetermined structured checklists and the involvement of examiners through the implementation process ensured objectivity which is the main tenet of OSCE (Wilkinson et al., 2003).

Lessons Learnt

Preparation and Orientation of Students

Important lessons were drawn from the first experience of conducting OSCE in the Department of Nursing Sciences. It was clear at the end of the examination that clarity of purpose of OSCE is of utmost importance for examiners and especially students considering that they were switching from the traditional kind of assessment. Being the first time students were being assessed through an OSCE, it was very stressful for students. This was evidenced by the questions that were asked by the students during the orientation meeting. Similarly, on the day of the examination, students exhibited stress as they awaited the assessment process. This could have affected their performance. As indicated earlier, studies have reported OSCE settings to be stressful or intimidating for participants although none has compared the level of stress to other forms of formal examinations (Evans, 2008; Rennie & Main, 2006). Orientation of students was also done late hence students felt ill prepared.

Cost of OSCE

OSCEs being costly require that ample time is allocated to the preparation process (Evans, 2008; Turner & Dankoski, 2008; Marwaha, 2011; Nulty et al., 2011). During the implementation, it was discovered that a lot of time, material and human resources were required to conduct effective OSCEs. This was compounded by a large number of students that were examined. Additionally, there were limited number of assessors and standardized patients. This could have affected the results and performance of students as some students were assessed outside the scheduled examination time. This could have resulted in fatigue in both the students and assessors thereby affecting the validity and reliability.

Weighting of the Questions/Scenarios

All the stations were allocated equal marks despite the fact that some skills on certain stations were more critical than others. This resulted in a situation where a student who fails to perform a critical skill at one station but performs well in a less critical skill, ends up passing the overall examination after aggregation of scores. This situation entails that a student may progress to the next level of training or graduates without that particular core competence.

Use of Checklists

Development of OSCE checklists as well as scoring is not as straightforward as it was thought. It appears that this also needs detailed discussion; attention and agreement especially between or among examiners assessing the same skill in cases where students are divided into more than one stream. Assessors should be adequately prepared to ensure consistence in approach and inter-rater reliability (Evans, 2008). It is therefore important that checklists are standardized and comprehensive thereby preventing disadvantaging students. It was discovered during the examination that some checklists used in the examination were not exhaustive. This could have introduced some bias as it prompted some assessors to begin probing for more answers from the students. Therefore, thorough preparation and pretesting of the checklists is required to avoid introducing bias and subjectivity.

Pre-Testing of OSCE

It is also recognized from our experience that it is important
to meticulously organize and pretest OSCE in order to uphold reliability and validity. The duration, interconnectedness, number and order of OSCE stations need to be carefully examined in order to ensure that the potentially competing requirements of validity and reliability are balanced because they all affect students (Rushforth, 2007).

Finally, it came to light that the process of OSCE becomes monotonous especially when examiners are dealing with huge numbers of students. At the same time, it is recognized that changing examiners for a particular station may compromise the objectivity of the assessment and disadvantage the students.

**Implications**

Implications for introducing OSCE in the Department of Nursing Sciences include:

OSCEs are very costly to implement. For example the number of faculty required to assess a group of 104 students using DOPS where each student is assessed on one procedure for 30 minutes would have been 13. For the OSCE where we had two streams running concurrently each with seven stations, the number was almost doubled. In addition to faculty we required 6 Standardized patients for each of the three observer stations in the two streams. Further each stream required a coordinator and time keeper. Apart from faculty there was also a cost for paying standardized patients. OSCEs are also time consuming in both preparations and actual administration. This cost implication entails a need to allocate adequate funds for the process thus demanding changes in student examination fees and changes in departmental examination budget.

Introducing OSCE will require investment in identifying and training simulated and or standardized patients well in advance of the assessment. It has been documented that, when well trained simulated and or standardized patients cannot be distinguished from real patients, are stable over time, and can provide accurate feedback in both preparations and actual administration. This cost implication entails a need to allocate adequate funds for the process thus demanding changes in student examination fees and changes in departmental examination budget.

Regarding students’ opinion of OSCE, a study conducted by Turner and Dankoski, to assess the validity, reliability and feasibility, majority of students felt that they had been fairly tested (Newbel, 2004). Despite the above challenges, Newble (2004), recognized that traditional clinical examinations have serious limitations related to validity and reliability issues thus supporting the use of OSCE as an alternative emblem.

This paper also further draws from Newble’s (2004) assertion and agrees that OSCE is particularly suitable for assessing numerous components of clinical competence. However, attitudinal and behavioural aspects of the student may not be fully assessed through this method. Nevertheless, certain attitudinal issues may be assessed if well trained standardized patients are used. This therefore requires the use of other methods of assessment that are best suited for assessing behavioural and attitudinal aspects of candidates such as DOPS in the actual clinical environment.

Finally OSCE must be integrated into the overall clinical evaluation system to be used for both formative and summative assessment. Use of OSCE for formative assessment, allows for provision of immediate feedback and serves as a teaching opportunity (Newbel, 2004).

**Discussion**

OSCE as a performance-based assessment method is a well established student assessment tool for many reasons: competence-based, valid, practical and effective means of assessing clinical skills that are fundamental to the practice of Nursing and other health care related professions (Association of American Medical Colleges, 2008; El Darir & Abd El Hamid, 2013). Its popularity as a major tool in the assessment of clinical competence is well documented as being specifically prominent in assessing situations where reliability and content validity are fundamental elements for making the results of such assessments justifiable to both examinees and external agencies (Newbel, 2004; Rushforth, 2007; Selim et al., 2011; El Darir & Abd El Hamid, 2013). The popularity of OSCE resulted from concerns that were raised about the traditional clinical and oral examinations used for assessing clinical competence (Rushforth, 2007; Holmboe & Hawkins, 2008; Levette-Jones et al, 2010; Marwaha, 2011; El Darir & Abd El Hamid 2013). The concerns were triggered by the discovery of low correlations between examiners marks allocation, which resulted in unacceptable reliability. However, change in some parts of the world took long to occur due to a general lack of an optional assessment method for clinical competence.

The introduction of OSCE has provided benefits in clinical testing to that of objective written examinations in knowledge testing. The use of marking which is based on a checklist has lead to improved inter-rater consistency (Rushforth, 2007). Testing students’ performance on numerous stations has also contributed to the increase in the number and range of competencies that could be sampled thus improving on content validity (Downing, 2003; Bhat & Anal, 2006). It is these benefits of OSCE that outweighs its cost (Selim et al., 2011; El Darir & Abd El Hamid, 2013).

Regarding students’ opinion of OSCE, a study conducted by Turner and Dankoski, to assess the validity, reliability and feasibility, majority of students felt that they had been fairly marked (Turner & Dankoski, 2008). Similarly, OSCEs are regarded by most students as comprehensive-covering a range of knowledge and clinical competences and useful practical experience (Piere, 2004). Similar assertions were made by Eldarir and Abd el Hamid (2013) where students reported to have had positive opinion of OSCE when they isolated a number of advantages of OSCE compared to traditional evaluation; measuring of course objectives, enhancing teaching level, relating theory to practice, making examinations well developed, increasing decision making abilities and an enhanced method of evaluation (El Darir & Abd El Hamid, 2013). On the other hand negative student opinions on OSCE have also been reported for example in a study by Moudoon, Biesty and Smith (2013), 57% students either disagreed or strongly disagreed to the statement that OSCE reflected real life clinical situation.

**Limitations of the Study**

There are two main limitations of this case study: Firstly OSCE was being introduced for the first time and used for summative assessment of clinical competencies. This could have affected both students and examiners in some unique way as there was no prior experience with this type of testing during
formative assessments. For students, introduction of a new assessment technique could have increased the levels of anxiety in addition to the usual anxiety associated with examination consequently affecting performance. Therefore, the obtained performance level could not have reflected the actual levels of competencies in the tested clinical skills. For examiner, switching from DOPS to OSCE could have affected inter-rater reliability. Despite the orientation some examiners could not confine to the structured checklist, they still asked additional questions to some and not all students as is the case in DOPS. Second limitation was the use of untrained SPs. In cases where SPs were required, postgraduate students were used. Although the SPs were oriented to their role, the orientation was not adequate such that they varied their responses in some cases, thus infringing on standardization which is a critical component in ensuring objectivity in OSCE.

Conclusion

OSCEs are suitable for testing clinical, technical and practical skills which may not be adequately assessed through traditional assessment methods, and it possesses the ability to improve the validity and reliability of assessments of many aspects of clinical competence. It should therefore be used as a method for assessing nursing students’ clinical competences, however careful consideration should be taken to avoid entirely relying on OSCE as the only means of assessing clinical competence. Hence, the use of other methods of assessment that are best suited for other aspects of clinical competence may be equally important in complementing OSCE. In addition we acknowledge that while OSCE as an assessment tool has been widely researched and documented, research evidence specific to nursing in Zambia is scanty. This therefore underscores the critical need for more research in this area in Zambia.

Acknowledgements

The authors acknowledge the Department of Nursing Sciences at the School of Medicine, University of Zambia for facilitating the Implementation of OSCE.

REFERENCES


Muldoon, K., Biesty, L., & Smith, V. (2013). I found the OSCE very stressful: Student Midwives attitudes towards an objective, structured clinical examination (OSCE) nurse education today.


