Evaluation of the Nursing Students' Skills by DOPS

Mohammad Jalili, MD[1]; Masoomeh Imanipour, MSc[2*;] Nahid Dehghan Nayeri, PhD[2];

Azim Mirzazadeh, MD[1]

[1] School of Medicine; Tehran University of Medical Sciences, Tehran, Iran
[2] Nursing and Midwifery Care Research Center (NMCRC); School of Nursing and Midwifery; Tehran University of Medical Sciences, Tehran, Iran

Abstract

Background and purpose: *Student assessment is an important issue in medical education. It is necessary to use objective, valid and reliable assessment methods that are appropriate to the learning objectives and domain.The objectives of this study were 1- Development of the DOPS test (Direct Observation Procedural Skills) and identification of its validity, reliability and acceptability. 2- Evaluation of the nursing students' skills by the DOPS test.*

Methods: *First, the DOPS test was developed by reviewing the literatures and other related tools. Its validity was determined by computing the Content Validity Ratio (CVR) and Content Validity Index (CVI), based on the experts' panel opinions. Then, the clinical skills of the students in doing special procedures were evaluated by the DOPS test. The Cronbach Alpha coefficient was used to determine the DOPS reliability. To explore the satisfaction with the DOPS test, a questionnaire and VAS were used.*

Results: *Based on the results, the DOPS test showed high content validity with CVR=1 and CVI=0.94. The reliability of the DOPS test confirmed α>0.7. The respondents held a positive attitude toward the DOPS test with mean satisfaction greater than 8. The performance of the majority of students in doing special procedures was good based on the DOPS scores.*

Conclusions: *The DOPS test is a suitable evaluation method to assess the students' procedural skills objectively because of its high validity, reliability and acceptability. To evaluate all aspects of the students' performance the DOPS test should be combined with other clinical performance tests.*

Keywords: *STUDENT ASSESSMENT, CLINICAL ASSESSMENT, VALIDITY, RELIABILITY*

Introduction

Evaluation is an inseparable part of the education process and student assessment is one of the main subjects of evaluation among the bases of academic education (1, 2). Assessment can reinforce the students' intrinsic motivation to learn and inspire them to set higher standards for themselves (3). Besides, the public needs to be assured that the graduates have the necessary professional qualifications and the required skills for performing their professional duties (1, 4). Assessment is a method of protecting the public by identifying incompetent physicians (4). With respect to this issue, as well as due to the complexity of the concepts, the plurality and the diversity of educational objectives and the critical tasks that students ought to perform in their profession, student assessment plays a major role in the course of medical education (1).

Over the past decade, medical schools have taken great and new efforts to provide accurate and reliable methods to assess the knowledge, behavior and professional characteristics of the medical students (4). In selecting the appropriate assessment instruments, it is essential to consider what educational objectives and learning domains

∗**Corresponding author:** Masoomeh Imanipour, MSc. Nursing and Midwifery Care Research Center (NMCRC); School of Nursing and Midwifery; Tehran University of Medical Sciences, Tehran, Iran.
Email: m_imanipour@tums.ac.ir

are. In this vein, Ven der Veluten (2002) added some efficient assessment techniques for each of stage of the Miller's level pyramid. The Learning Levels of Miller's pyramid and the suggested assessment tools for each one _from bottom to top_ are: "Know" (written questions), "Know How" (written questions), "Show How" (OSCE), "Does" (Portfolio, Logbook, Peer review, DOPS) (1, 5). Hence, it is obvious that work-based assessments or Performance tests are best suited to assess the highest level of the pyramid - "Does" (5, 6). In other words, to gather data regarding the students' performance in its real practice, we need performance tests (7). Performance tests evaluate the learning processes and products directly. In fact, these types of tests deal with skills and assess the educational objectives in a direct manner (7-9). Currently, there are various performance tests and work-based assessments that each, regarding its limitations and advantages, can be applied in different fields. Direct Observation of Procedural Skills (DOPS) is a performance test which focuses on evaluating the "Does" field by observing the trainee in the real workplace setting. The DOPS test is a structured rating scale for assessing and providing feedback on practical procedures. The student is observed directly by an expert examiner in a few real encounters that last for about 15-25 minutes. Direct observation of the procedural skills is practical and easy to use. This technique permits feedback based on the students' real performances; hence, it can be used as a method for improving the students' performance. Trainees generally welcome the opportunity to be observed by someone who is more experienced than them and it gives them immediate feedback (5, 10, 11). The DOPS test has been mainly used to assess the procedural skills of the graduates of medicine or surgical residents (12, 13), and simultaneously it can be applied to nursing or any other health care careers (10).

As one of the most important goals of nursing education is the improvement of the practical skills of the nursing students in order to equip them with the qualities and qualifications of a graduate nurse, it is imperative to evaluate the student performance with valid, reliable, and objective scales like DOPS, OSCE and portfolio (14). Despite this, the findings of several studies have shown that one of the major problems in the clinical assessment of nursing students is the lack of consistency, because many traditional strategies of clinical assessment are based on the unstructured and experimental observation of student performance by the instructor, which runs the risk of observer bias. Therefore, the quality of students' assessment by the professors has been permanently viewed with skepticism (15-18). For example, students participating in the qualitative study of Calman et al., (2002) had little confidence in the methods that were utilized to assess their clinical competence. They claimed that insufficient attention was paid to practical skills and they were also concerned regarding the way their major nursing skills were assessed (15).

With respect to these problems and according to the complexity of the technical nursing skills in Intensive Care Units, the assessment of the students' performance in these clinical settings requires an objective work-based assessment. The aim of this study is to evaluate the nursing students' performance in performing some common procedures of the Intensive Care Units by a performance test termed DOPS. Also, the paper describes the process of developing the DOPS test and determines its validity, reliability and acceptability.

Methods

This is a cross-sectional study to evaluate the nursing students' procedural skills in the ICU and Intensive Care Units open heart (ICUOH). This study was conducted in the hospitals of the Tehran University of Medical Sciences. The subjects of the study were the nursing students in their internship course, during the seventh educational semester. The sample size was 38, which consisted of all the students in the seventh semester. The procedural skills were assessed by DOPS. To achieve this, special procedures that the

students were supposed to have learned in the critical care setting, were determined. These procedures consisted of gavage, CVP measurement, suctioning of ETT or tracheostomy and arterial line irrigation. The DOPS test was developed by reviewing the related literatures and other related tools. Its validity was determined by computing the Content Validity Ratio (CVR) and the Content Validity Index (CVI), based on the experts' ratings. Then, each student was evaluated by virtue of two different procedures employing the DOPS test, at least twice. The Cronbach Alpha Coefficient (method) was used to determine the reliability of the DOPS. We also identified the instructors and the students' satisfaction with the DOPS test, by a 0-10 centimeter Visual Analogue Scale (VAS).

Results

This study was analyzed in multi-dimensions. The CVR is an item statistic that is useful in rejecting or approving specific items in an assessment tool. To achieve this, all the expert panelists were asked to rate each of the items in a continuum of "essential", "useful but not essential" and "not necessary". The responses from all the panelists were pooled together and the number indicating "essential" for each item was determined. According to the CVR calculation formula and the Lawshe Minimal Values of the CVR Table, all the items of the DOPS tool were necessary and important at a significance level of P<0.05 (when the number of panelists are 13, CVR should be at least 0.54)

(Table 1).

After identifying the CVR and deciding whether the content domain is adequately represented by the items, the Item-Content Validity Index (I-CVI) and the Scale-Content Validity Index (S-CVI) were calculated. To accomplish this, the experts were asked to rate each item based on relevance, clarity and simplicity on a four-point scale. The result of the CVI calculation formula showed that the I-CVI of all the items was greater than 0.7 and the S-CVI was equal to 0.94 (Table 2).

The Alpha Cronbach showed high reliability of the DOPS test in evaluating the special procedures in both clinical settings (ICU, $\alpha=0.87$; ICUOH, $\alpha=0.92$) (Table 3).

The acceptability of the DOPS test and the satisfaction of the instructors and students with the DOPS test were determined with means of 8.87 and 8.18, respectively. None of the respondents were unsatisfied (Table 4). The DOPS test scores for evaluating the students' skills in performing special procedures were considerable. The majority of the students showed good performance and their scores were generally between 17 and 20. No failure scores (below 10) were recorded (Table 5).

Discussion

The evaluation of clinical competency is a complex and difficult task in medical education because many of educational objectives are psychomotor. There are many challenges in the assessment of the clinical performance of the students of medicine and other health-related disciplines. Commonly,

Table 1. Content Validity Ratio of the DOPS test

Item	Evaluated Character	CVR necessity
Item1	Preparation of patient, equipments, environment	1
Item2	Communicate with patient and explain the procedure	1
Item3	Clean or aseptic technique	1
Item4	Correct technique and meet standards	1
Item5	Correct sequence	1
Item6	Collect equipments and tidy patient	1
	Scale Content Validity Ration	1

Table 2. Content Validity Index of the DOPS test

Item	Evaluated Character	CVI relevancy	CVI simplicity	CVI clarity
Item1	Preparation of patient, equipments, environment	1	1	0.92
Item2	Communicate with patient and explain the procedure	1	1	1
Item3	Clean or aseptic technique	1	1	1
Item4	Correct technique	1	0.85	0.92
Item5	Correct sequence	1	0.92	0.92
Item6	Collect equipments and tidy patient	0.92	1	1
Scale Content Validity Index		0.94		

Table 3. Reliability of the DOPS test

Procedure	Clinical Setting ICU	Clinical Setting ICUOH
Suctioning	0.82	----
Gavage	0.83	----
Arterial line irrigation	----	0.82
CVP measurement	----	0.92
Scale Reliability	0.87	0.92

the traditional student assessment methods are based on unstructured teachers' observation, oral tests and teachers' opinions, which produce unreliable, invalid and subjective evaluations. Dissatisfaction with such assessment methods and improvement in the evaluation techniques have led to new efforts to provide accurate, reliable and valid assessments of students' competencies (17). The OSCE, mini-CEX, DOPS are some of the assessment methods for performance evaluation in the psychomotor educational domains.

In this study, we applied the DOPS test to evaluate the procedural skills of the nursing students according to the agreement of all the expert panelists. The DOPS test is suitable for the evaluation of the "Does" level of Miller's pyramid (5, 10). Several authors comment on the lack of rigorous testing of the procedural skills, but DOPS is an instrument typically used to assess the procedural skills of the doctors at any level (18). The DOPS test is one of the new assessments being piloted in the UK as part of the new "Foundation Program" for medical graduates in the first two years (12). Also, the Australian College of Physicians has applied the DOPS test as part of the Professional Standards Program since 1994 (13). DOPS is an instrument used most prevalently to test the surgical residents, but in recent years some authors have suggested that it can also be applied to the junior doctors to assess their general performance and that of Nursing, Midwifery and other disciplines (10, 12). All the methods of assessment have strengths as well as intrinsic flaws. While selecting a method, some features need to be considered. Van der Vleuten describes five criteria for determining the usefulness of a particular method of assessment: reliability, validity, impact on learning, acceptability to learners and faculty and costs (1, 4, 19). Content validity, as one of the most important criteria, is representative of the content and depends on the adequacy of a specified domain of

Table 4. Satisfaction with the DOPS test

Group	Maximum Score	Minimum Score	mean±SD	Satisfaction Unsatisfied 0-3	Satisfaction Moderate Satisfaction 4-6	Satisfaction Satisfied 7-10
Instructors	10	5	8.87±1.64	0%	12.5%	87.5%
Students	10	4	8.18±1.62	0%	21.6%	78.94%

Table 5. Evaluation of students' performance in doing special procedures

Clinical Setting	Procedure	mean±SD	Student Performance			
			Unaccepted 0-9	Poor 10-13	Moderate 14-16	Good 17-20
ICU	Suctioning	17.78±1.5	0%	2.6%	26.3%	71.1%
	Gavage	18.54±1.4	0%	0%	23.7%	76.3%
ICUOH	Arterial line irrigation	18.43±1.6	0%	0%	18.4%	81.6%
	CVP measurement	17.94±2.2	0%	0%	21.1%	68.4%

content that is sampled (20). When a new scale is developed, researchers are expected to provide extensive information regarding its validity and reliability. The content validity of an instrument is obtained from some sources. The CVR and CVI are quantitative approaches to determine the content validity and nursing researchers use this method most extensively (21). In this study we too applied this approach. According to the results and based on the Lawshe minimal values of CVR, when expert panelists are 13, the minimum acceptable CVR is 0.54 (22). Table 1 shows that all the items in the DOPS test in this study had CVR=1, and that all the items were essential to evaluate the special procedural skills of the nursing students in the critical care settings at a significance level of $p<0.05$. After the items were identified for inclusion in the final form, the content validity index was computed for each item and scale totally. Researchers use the I-CVI information to guide them in revising, deleting or substituting the items (21). The items that have CVI over 0.79 remained unchanged, whereas those with CVI=0.79-0.70 need to be revised, and the ones with CVI less than 0.70 need to be omitted (23). Based on these findings, the I-CVI was over 0.79 for all of the items in this study. Many authors have indicated that an S-CVI of 0.80 or more is acceptable (24). This was equal to 0.94 and the validity of the DOPS test was confirmed as well.

The Reliability of an instrument is another useful criterion. Reliability refers to the accuracy of the results, which implies results stability, consistency and reproducibility (8). As the second objective to determine the DOPS reliability, we applied the test to each of the students and then calculated the Cronbach Coefficient. Results showed $\alpha>0.80$ for all procedures, which confirms a high reliability.

Although there is little evidence regarding the design and the application of DOPS and further studies are warranted, the Royal College of Physicians, which has developed a DOPS instrument for the Foundation Program, anticipates that it will be found to be a highly valid and reliable instrument (25, 26).

Van der Vleuten believes the stakeholders' acceptability of an instrument is a criterion for good assessment (27). Researches regarding the acceptability of DOPS are rather few in number; however, they appear to be acceptable to both the learners and examiners (26). There are several methods of reviewing the acceptability of an instrument from the perspective of both the examinees and the examiners, like the satisfaction dimension. In the present study, we examined the instructors' and the students' satisfaction levels with the DOPS test by employing the Visual Analogue Scale (VAS). The mean score of satisfaction in both groups was more than 8 and the majority of stakeholders were satisfied with this clinical assessment method.

Scores of the DOPS tests that were performed to evaluate the students' performance in performing special procedures demonstrated good performance of the students with a mean score above 17 (A score) for all the procedures. There were no points of failure in four procedures.

Based on the findings, the DOPS test can be applied as an objective assessment method to evaluate the students' procedural skills.

However, further studies on the other procedures and different clinical settings are warranted. Although the results illustrated good performance of the students in performing the procedures, for a more accurate and precise assessment of clinical competency it is suggested that the DOPS test be blended with the other clinical assessment techniques. As a matter of fact, multiple informational resources are necessary to be able to cover all the aspects of clinical competencies.

Acknowledgment

References

1. Shumway James M, Harden RM. The Assessment of Learning Outcomes for the Competent and Reflective Physician. AMEE Guide No. 25. Medical Teacher. 2003;25(6):569-84.
2. Kimeili Gh, Rezaei Gh. Students Assessment by Faculty of Basic Sciences in Zahedan Medical University Of Sciences. Medical Education Journal. 2002;4:49-53 [Persian].
3. Friedman Ben-David M. The Role of Assessment in Expanding Professional Horizons. Medical Teacher. 2000;22:472-7.
4. Epstein Ronald M. Assessment in Medical Education. The New England Journal of Medicine. 2007;356:387-96.
5. Zubeir A, Chong Yap S, Khoo Hoon E. Practical Guide to Medical Student Assessment. Singapore: World Scientific. 2006.
6. Norcini John. ABC of Learning and Teaching in Medicine. Work Based Assessment. British Medical Journal. 2003;326:753-5.
7. Nitko Anthony J. Educational Assessment of Students. 3rd ed. New Jersey: Merrill Prentice Hall. 2001

8. Seif AA. Educational Measurement, Assessment and Evaluation. 3rd ed. Tehran: Doran Publisher; 2004. [Persian]
9. Kubiszyn T, Borich G. Educational Testing and Measurement, Classroom Application and Practice. 7th ed. United States: John Wiley & Sons, INC. 2003
10. Swandwich T. Understanding Medical Education: Evidence, Theory and Practice. 1st ed. Malaysia: Wiley-Blackwell. 2010
11. Review of work-based assessment methods. Centre for Innovation in Professional Health Education & Research (CIPHER).The University of Sydney, Australia. 2007.
12. Beard J, Strachan A, Davies H. Developing an education and assessment framework for the Foundation Program. Medical Education. 2005;39(8):841-51.
13. Newble D, Paget N, McLaren B. Revalidation in Australia and New Zealand: approach of the Royal Australasian College of Physician. BMJ 1999;319:1185-8.
14. Karayurt O, Mert H, Beser A. A Study on Development of a Scale to Assess Nursing Students' Performance in Clinical Settings. Journal of Clinical Nursing. 2008;18:1123–30.
15. Calman L, Watson R, Norman I, Refern S, Murrells T. Assessing Practice of Student Nurses: Methods, Preparation of Assessors and Student Views". Journal of Advanced Nursing. 2002;38(5):516–23.
16. Dolan Gina. Assessing Student Nurse Clinical Competency: Will We Ever Get It Right?" Journal of Clinical Nursing. 2003;12(1):132-41.
17. Harden RM, Stevenson M, Downie W. Assessment of Clinical Competence Using Objective Structured Clinical Examination. British Medical Journal. 1975;1:447-51.
18. Sidhu R, Grober E, Musselman L, Reznick R. Assessing competency in surgery: Where to begin? Surgery. 2004;135(1):6-20.
19. Tahernejad K, Javian F. Advanced Clinical Performance Assessment of Medicine Students: Challenges, Methods and Approaches". Strides in Development of Medical Education. 2009;5(1):58-70. [Persian]
20. Yaghmaie F. Content Validity and Its Estimation". Journal of Medical Education. 2003;3(1):25-7.
21. Polit Denise F, Cheryl Tatano B. The Content Validity Index: Are You Sure You Know What's Being Reported? Critique and Recommendations". Research in Nursing & Health. 2006;29:489–97.

22. Lawshe CH. A Quantitative Approach to Content Validity". Personal Psychology. 1975;28:563-75.

23. Lynn MR. Determination and Quantification of Content Validity. Nursing Research. 1986;35:382–5.

24. Davis LL. Instrument review: Getting the most from your panel of experts. Applied Nursing Research. 1992;5:194-7.

25. Wilkinson J, Benjamin A, Wade W. Assessing the performance of doctors in training. BMJ. 2003;327:91-2.

26. Davies H, Archer J, Heard S. Assessment tools for Foundation Programs -a practical guide. BMJ Career Focus. 2005;330(7484):195-6.

27. Norcini J, Anderson B, Bollela V, Burch V, Joaˆ O Costa M, Duvivier R, Galbraith R, Hays R, Kent A, Perrott V, Roberts T. Criteria for Good Assessment: Consensus Statement and Recommendations from the Ottawa 2010 Conference. Medical Teacher. 2011;33:206–14.