



Direct observation of procedural skills (DOPS) evaluation method: Systematic review of evidence

Masoumeh Erfani Khanghahi¹, Farbod Ebadi Fard Azar²

Received: 23 Jan 2017

Published: 3 June 2018

Abstract

Background: Evaluation is one of the most important aspects of medical education. Thus, new methods of effective evaluation are required in this area, and direct observation of procedural skills (DOPS) is one of these methods. This study was conducted to systematically review the evidence involved in this type of assessment to allow the effective use of this method.

Methods: Data were collected searching such keywords as evaluation, assessment, medical education, and direct observation of procedural skills (DOPS) on Google Scholar, PubMed, Science Direct, SID, Medlib and Google and by searching unpublished sources (Gray literature) and selected references (reference of reference).

Results: Of 236 papers, 28 were studied. Satisfaction with DOPS method was found to be moderate. The major strengths of this evaluation method are as follow: providing feedback to the participants and promoting independence and practical skills during assessment. However, stressful evaluation, time limitation for participants, and bias between assessors are the main drawbacks of this method. Positive impact of DOPS method on improving student performance has been noted in most studies. The results showed that the validity and reliability of DOPS are relatively acceptable. Performance of participants using DOPS was relatively satisfactory. However, not providing necessary trainings on how to take DOPS test, not providing essential feedback to participants, and insufficient time for the test are the major drawbacks of the DOPS tests.

Conclusion: According to the results of this study, DOPS tests can be applied as a valuable and effective evaluation method in medical education. However, more attention should be paid to the quality of these tests.

Keywords: Evaluation, Medical education, Directly observed procedural skills

Copyright© Iran University of Medical Sciences

Cite this article as: Erfani Khanghahi M, Ebadi Fard Azar F. Direct observation of procedural skills (DOPS) evaluation method: Systematic review of evidence. *Med J Islam Repub Iran.* 2018(3 June);32:45. <https://doi.org/10.14196/mjiri.32.45>

Introduction

It has been proven that a country's cultural, social, political, and economic success relies on a coherent and dynamic educational system. Only with such a system, a county can keep pace with social and industrial developments and confirm a high position among successful global countries. Education and training courses cannot help the system in achieving its goals, as training should be in accordance with scientific principles and methods to meet the needs involved. Otherwise, trainings will be fruitless and in some cases they may even lead to the waste of capital in a system (1-3).

One question that needs to be raised, however, is whether

the results of the training courses are in line with the set goals or not. Most scholars have explicitly answered this question: "A complete and comprehensive evaluation could make us aware of the effectiveness of the training outcomes and provide a feedback assessment." (4, 5) Training medical students is one of the most important areas in education. The goal of medical education is training physicians to enable them to address the essential needs of the society in addition to passing written tests (6, 7). The written tests merely examine the initial level of Miller's Pyramid, which is the assessment of clinical competence, thus, the actual performance of students, which is similar to their future

Corresponding author: Dr Masoumeh Erfani Khanghahi, afsaneerfani@gmail.com

¹ Department of Educational Management, Faculty of Management and Accounting, Yadegar-e-Imam Khomeini (IAH), Shahre Rey Branch, Islamic Azad University, Tehran, Iran.

² Health Promotion Research Center, Iran University of Medical Sciences, Tehran, Iran.

↑What is "already known" in this topic:

Direct observation of procedural skills (DOPS) was acceptable in 3 aspects as follows: (1) validity and reliability of DOPS, (2) satisfaction with DOPS, and (3) comparing the use of DOPS and mini- CEX with traditional methods.

→What this article adds:

This study presented strengths and drawbacks of DOPS. Moreover, in this study training impacts of DOPS, performance of participants using DOPS tests, and quality of conducting DOPS tests were examined.

professional status at the high levels of the pyramid, is not assessed. On the other hand, the final evaluation cannot assess the learners' performance (8, 9). Hence, the teachers tend to use formative assessment to become aware of the students' progress. This is an attempt to ensure high quality in educational programs and is used to motivate students towards what they need to learn. Consequently, education will result in outcome-based education (8, 10). In recent years, various evaluation methods have been used in medical education (11-15). One of the most important and well-known of which is direct observation by students during the practical skills. Direct observation of procedural skills (DOPS) method is considered as one of the most well-known models of this type of evaluation (16, 17). This method was specifically designed to evaluate practical skills and provide feedback; it requires direct observation of an assistant during a procedure and coincides with evaluation in a written form. This method is particularly useful in evaluating the practical skills of the assistant objectively and systematically (18). In this method, observation of the assessor is documented in a checklist, and then the trainee is provided with a feedback based on objective findings. The number of tests varies depending on the basic skills required for learning, which may include up to 8 tests during a period (19-21). The method gives the trainees an opportunity to receive constructive feedback and directs their attention to the essential skills required to perform the procedure on the grounds that the evaluation is aimed to improve the performance and ask for a specific and on time feedback (22). Reviewing the literature on modern methods of evaluation including DOPS indicates that despite a considerable amount of studies based on these methods, the use of modern methods of evaluation has not yet been taken in to account efficiently. Thus, given the importance of medical education and the effectiveness of new methods of evaluation, further research and development of new methods of evaluation is required. This paper was conducted to systematically review the studies on DOPS and provide useful information for the development and use of this evaluation method in medical training.

Methods

Search strategy

This was a systematic review developed and implemented in 2016. Data were collected searching such keywords as evaluation, assessment, medical education, validity, reliability, content validity ratio, content validity index, direct observation of procedural skills, and DOPS in Persian and English languages on Google scholar, PubMed, Science Direct, and SID search engines and by searching unpublished data (Gray literature) and selected references (reference of reference). The validity and reliability of DOPS were also examined in this study.

Inclusion and exclusion criteria

No time limitation was considered in searching for papers, and published papers only in Persian and English were searched. Studies conducted on medical education were included in this review, and exclusion criterion was the concurrent effect of DOPS with other evaluation methods.

Data extraction and quality assessment

To evaluate the quality of the extracted papers, 2 assessors evaluated them based on the checklists involved. First, the titles of all papers were reviewed and the papers incompatible with the objectives of the study were excluded. Subsequently, abstracts and full texts of papers were studied and those that were least relevant to the objectives of the study were identified and excluded.

After evaluations by the checklist, which were supervised by the reviewing team, papers from our checklist included the following information: (1) name of the first author, (2) years of publication, (3) country, (4) participants, and (5) findings. The selected papers were fully studied and evaluated and the required results were extracted and summarized based on the designed tables (extraction table). End-NoteX5 software was used to organize the study and identify the frequency cases (Fig. 1).

Results

The initial search result entailed 236 papers. Once non-

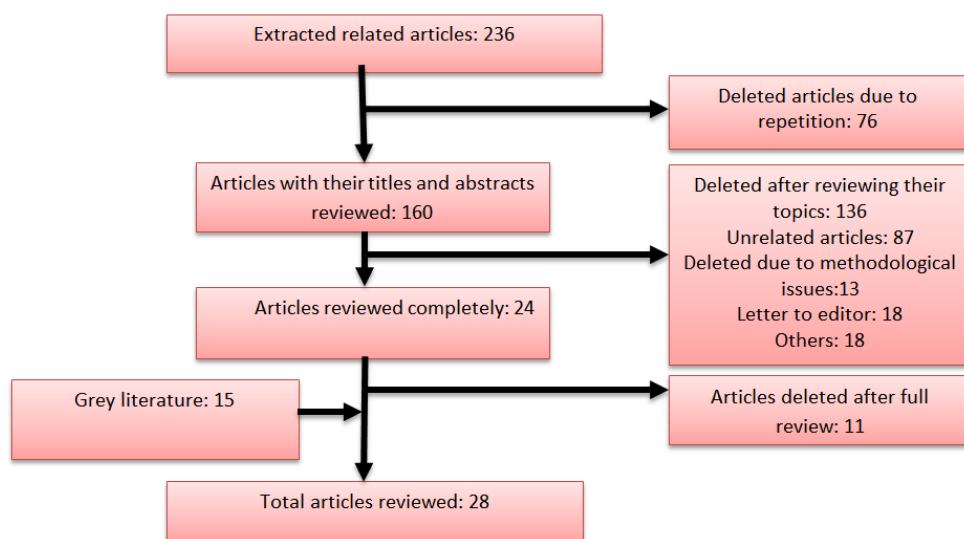


Fig. 1. Articles selection and search process

related papers and those shared in various databases were excluded, 28 papers were selected. The results are summarized in Tables 1 through 6. After classifying and analyzing the results of the study, we presented the outcomes in the following 7 areas: satisfaction with DOPS, strengths and drawbacks of DOPS, impact of DOPS assessment, reliability and validity of DOPS, performance of participants in DOPS, and quality of holding evaluation sessions with respect to DOPS.

Satisfaction with DOPS

The results of participants and assessors' satisfaction

with DOPS were reported in 5 studies and indicated moderate satisfaction with DOPS (Table 1).

Strengths and Drawbacks of DOPS

The perspectives of participants and assessors about DOPS test were reported in 8 studies. In the reviewed studies, 27 strengths and 16 drawbacks were noted for this method of assessment (Table 2). The major strengths of this evaluation method mentioned in the studies included providing a feedback to participants, promoting practical skills of participants, independence during

Table 1. Satisfaction of trainees and assessors from direct observation of procedural skills (DOPS) assessment method

Author: year	Country	Participants	Findings
Asadi K, et al., 2012(23)	Iran	70 orthopedic interns	About 57% of participants were satisfied with performing DOPS evaluation. Minimum level of satisfaction was related to devoting time for each procedure and noncompliance of patients with students.
Hoseini L, et al., 2013(24)	Iran	67 undergraduate midwifery students	Students in DOPS group were significantly more satisfied than those using the current method.
Kang Y, et al., 2009(25)	USA	62 students who rotated on medicine clerkship	Rate of satisfaction of students and observers were Moderate or high
khoshrang H, et al., 2011(26)	Iran	57 residents (anesthesiology, surgery, urology, ENT, neck surgery, and neurology)	The students' satisfaction score was 41.40 ±5.23. DOPS assessment had a low to Moderate satisfaction rate among residents
Farajpour A, et al., 2012(27)	Iran	60 medical internship students of the emergency ward	Average expected score for satisfaction was set higher than 50. Satisfaction rate score was 76.7±11.6.

Table 2. Pros and cons of DOPS assessment method

Author: year	Country	Participants	Advantage	Disadvantage
Cohen SN, et al., 2009(28)	UK	138 dermatology specialists	<ul style="list-style-type: none"> ➤ Feedback or constructive criticism ➤ Supervision or observation ➤ Reassuring ➤ Good to be assessed on surgical skills 	<ul style="list-style-type: none"> ➤ Time-consuming to do ➤ Difficult to organize with a consultant ➤ Stressful or artificial ➤ Disagreement over correct technique ➤ Difficult to find appropriate cases
Wiles C, et al., 2007	UK	27 trainees working in an NHS neurology department	<ul style="list-style-type: none"> ➤ Observation and feedback 	-
Bindal N, et al., 2013(29)	UK	90 trainees and 129 assessors in anesthetic training program	<ul style="list-style-type: none"> ➤ Useful method of assessment 	<ul style="list-style-type: none"> ➤ Not helpful for training ➤ DOPS are a tick box exercise ➤ Not a proof of competency ➤ The need for training in DOPS
Wilkinson J, et al., 2008(30)	UK	177 medical specialists	<ul style="list-style-type: none"> ➤ Provided useful basis for feedback discussion ➤ Method said to be valid ➤ Value of formalized assessment process ➤ Improved training as a result ➤ High motivation to learn 	-
Cobb K, et al., 2013(31)	Netherlands	70 final year veterinary students	<ul style="list-style-type: none"> ➤ Prompted a deeper and higher reflective approach ➤ Acceptable method for practical skills assessment ➤ More opportunity for feedback 	<ul style="list-style-type: none"> ➤ Conflict between learning DOPS assessment and competent practitioner learning ➤ Increased stress levels
Dabhadkar S, et al., 2014(32)	India	7 students of Second year postgraduate OBGY	<ul style="list-style-type: none"> ➤ Opportunity to verbal and written feedback from the assessor ➤ High relevance to curriculum ➤ Acceptability ➤ DOPS can lead to better clinical management in long run 	-
Amini A, et al., 2015(33)	Iran	7 orthopedic residents and 9 faculty members	<ul style="list-style-type: none"> ➤ Opportunity for assessing clinical skills ➤ Self-assessing practical skills ➤ Pros and cons recognizing by students ➤ Students independence at the time of assessment ➤ Students become prepared for final exam ➤ More opportunity to communicate with faculty members 	<ul style="list-style-type: none"> ➤ Variation between the quality of assessment by assessor ➤ Bias by assessor ➤ Increased stress levels and confounding situation
Akbari M and Mahavelati Shamsabadi R, 2013(34)	Iran	110 dentistry students	<ul style="list-style-type: none"> ➤ Students independence at the time of assessment ➤ Improved training as a result 	<ul style="list-style-type: none"> ➤ Bias by assessor ➤ Stressful or artificial

Table 3. Educational impact of DOPS assessment method

Author: year	Country	Participants	Intervention	Results and conclusion
Tsui K, et al., 2013(35)	Taiwan	Students	Validity of students' measurement of prostate volume in predicting treatment outcomes	DOPS assessment improved students' prostate volume measurement skills (Cronbach's $\alpha > 0.70$).
Profanter C and Perathoner A, 2015(36)	Austria	193 fourth year students	Prospective randomized trial (DOPS vs. classical tutor system); surgical skills-lab course	DOPS group has high level of clinical skills. DOPS dimensions seem to improve tutoring and rates of performance.
Scott DJ, et al., 2000(37)	USA	22 junior surgery residents	2 weeks of formal video training	Trained group had significantly better performance than control group based on the assessment through direct observation ($P = 0.02$) compared to video tape assessment (NS). DOPS showed improved performance of participants after formal skills training on a video trainer.
Shahgheibi Sh et al., 2009(38)	Iran	73 medical students (42 control & 31 intervention)	Evaluation of the effects of DOPS on clinical externship of students learning level in obstetrics ward	DOPS group had significantly improved their skills than control group ($p = 0.001$). DOPS can be more useful in improving students' skill.
Morris A, et al., 2006(39)	UK	27 preregistration house officer	Perceptions of preregistration house officer	Participants mentioned that DOPS can improve their clinical skills as well as their future careers.
Hengameh H, et al., 2015 (40)	Iran	Nursing students	Comparing the effect of DOPS and routine assessment method	DOPS has significantly improved nursing students' clinical skills ($p = 0.000$).
Nazari R et a., 2013(41)	Iran	39 nursing students	Comparing the effect of DOPS and routine assessment method	DOPS has significantly improved nursing students' clinical skills ($p < 0.001$).
Dabhadkar S, et al., 2014(32)	India	7 second year postgraduate students of OBGY	Assessment impact of DOPS on students' learning	Five of 6 students who performed unsatisfactorily in the first round of DOPS moved to satisfactory level of performance in the second round of DOPS. Participants showed improvement in the second round of DOPS.
Amini A, et al., 2015(33)	Iran	7 orthopedic residents and 9 faculty members	Assessment impact of DOPS on students' learning	Students' performance was improved in the second stage of DOPS (from 50.6% to 59.4%). DOPS assessment methods had an effective role in increased level of students' learning and skills.
Bagheri M, et al., 2014(42)	Iran	Emergency medicine students (25 in experiment and 21 in control group)	Assessment impact of DOPS on students' learning	Experimental group had significantly high mean scores compared to the control group ($p = 0.0001$, $t = 4.9$).

assessment, great relevance to the courses and skills required, acceptability of this approach by participants, and its formative nature. However, the drawbacks are stressful evaluation, time limitation, and bias between assessors, and thus a great deal of coordination is required to employ the method.

Training impacts of DOPS

The impact of evaluation based on DOPS was reported in 10 papers (Table 3). In all studies, it was reported that DOPS had a positive effect on the performance of students. In studies that evaluated the effect of DOPS compared to conventional methods, DOPS was found to be more effective than other methods. In addition, several studies have shown that the performance of students improved after the

first stage of evaluation using DOPS.

Reliability and validity of DOPS

Reliability and validity of DOPS were assessed in 8 studies (Table 4). Content validity index and content validity ratio were used to determine the validity of DOPS, and the results revealed that the validity of DOPS is relatively acceptable. In addition, the opinion of assessors and participants about the validity of DOPS was sought in several studies. Most participants had confirmed the validity of DOPS. To examine the reliability of DOPS, interclass correlation coefficient, generalizability theory, and alpha coefficient were applied. Moreover, reliability of DOPS was confirmed in most studies. Feasibility of DOPS method was studied in a few studies.

Table 4. Validity, reliability, and feasibility of DOPS assessment method

Author year	Country	Participants	Validity	Reliability	Feasibility
Asadi K, et al., 2012(23)	Iran	70 orthopedic interns	CVI:0.90	0.80	-
Wilkinson J, et al., 2008(30)	UK	177 medical specialists	DOPS has low validity.	DOPS reliability can be favorably compared with the mini-CEX and MSF.	Mean time for observation in DOPS varied based on the procedure.
Watson MJ, et al., 2014(43)	Australia	Trainees in ultrasound-guided regional anesthesia (30 video-recorded)	'Total score' correlating with trainee experience ($r = 0.51$, $p = 0.004$)	Inter-rater: ICC = 0.25 internal consistency: ($r = 0.68$, $p < 0.001$)	The mean time taken to complete assessments was 6 minutes and 35 seconds

Table 4. Cntd

Hengameh H, et al., 2015 (40)	Iran	Nursing students	CVR: 0.62 CVI:0.79	Kappa coefficient:0.6 ICC: 0.5 G: 0.81	-
Barton JP, et al., 2012(44)	UK	157 senior endoscopists — 111 candidates and 42 assessors	Most of the candidates (73.6%) and assessors (88.1%) pointed out that DOPS assessment method was valid or very valid.		Scores of DOPS were highly correlated with assessment score of global experts.
Amini A, et al., 2015(33)	Iran	Seven orthopedic residents and 9 faculty members	CVI: 0.95	ICC:0.85	
Delfino AE, et al., 2013	UK	Six anesthesia staff for interviews, 10 anesthesiologists for consensus survey, and 31 anesthesia residents.	CVI: 0.9 kappa values: 0.8	G coefficient:0.90	
Sahebalzamani M, et al., 2012	Iran	55 nursing students	Correlation for theoretical: 0.117; correlation for clinical: 0.376	Cronbach alpha coefficient: 94 %	
Kuhpayehzade J, et al., 2014(45)	Iran	44 midwifery students	CVR: 0.75 CVI:0.50	Alpha coefficient: 0.81	

Table 5. Performance of trainees with direct observation of procedural skills (DOPS) assessment method

Author: year	Country	Participants	Results and conclusion
Bazrafkan L and et al., (46)-2009	Iran	54 dental students	Results showed that 86.7% of the students in various fields of dentistry had good performance and 13.3 % of the students had weak performance. In conclusion, DOPS is a suitable tool for assessing practical laboratory performance of dental students.
Mitchell C, et al., 2011(47)	UK	1646 trainees in a single UK postgraduate deanery	Statistical association was not found between scores of DOPS assessment methods and trainees' skills. DOPS mean scores are not suitable to predict lack of competence.
Dabhadkar S, et al., 2014(32)	India	Seven second year postgraduate students of OBGY	Six out of 7 students performed unsatisfactorily, and only one student had satisfactory performance.
Amini A, et al., 2015(33)	Iran	Seven orthopedic residents and 9 faculty members	According to the results, students had almost good performances (mean of good performance = 50.6%).

Performance of participants using DOPS test

Performance of participants was investigated by DOPS test in 4 studies and was found to be satisfactory in 2 studies, while participants did not have a good performance in the other 2 studies (Table 5).

Quality of conducting DOPS test

In 3 studies, the quality of conducting DOPS test was investigated from the perspective of both participants and assessors (Table 6). Not providing the necessary trainings for taking DOPS test, not being held at the scheduled time, not providing necessary feedbacks to participants, and insufficient time for the test are the major weaknesses of conducting DOPS tests.

Discussion

Study results revealed that satisfaction level of DOPS is

moderate. The strengths of this evaluation method mentioned in the studies included providing feedback to participants, promoting practical skills of participants, autonomy during evaluation, great relevance to the courses and required skills, acceptability of this approach by participants, and its formative nature. Some of the main weaknesses of this method are as follow: being stressful, the time limit for participants, bias/ dissimilarity of assessors, and requiring a great deal of coordination. Studies comparing the effect of DOPS and conventional methods revealed that DOPS is more effective than other methods. Several studies indicated that students' performance after the first stage of evaluation with DOPS was improved in the second stage. Furthermore, the results showed that the validity of DOPS is relatively acceptable, while its reliability and validity have been confirmed in most studies. Feasibility of DOPS method was investigated in few studies. Participants' performance who used DOPS was relatively satisfactory.

Table 6. Implementation quality of DOPS assessment method

Author: year	Country	Participants	Results and conclusion
Bindal N, et al., 2013(29)	UK	90 trainees and 129 assessors in anesthetic training program	About 33% of trainees and 43% of consultants were not receiving training about DOPS. DOPS assessments were not planned in many cases. Short time was spent on assessment. For most part of the assessment, feedback was achieved in 15 minutes.
Kang Y, et al., 2009(25)	USA	62 students who rotated on medical clerkship	About 77% of the observations were done while on call or during daily rounds. Furthermore, 73% of sessions were completed in 13 minutes or less. In 89% of sessions, students received verbal feedback at least for 5 minutes.
Morris A, et al., 2006(39)	UK	27 pre-registration house officers	Each individual DOPS was completed in less than 15 minutes. About 50% of trainees were aware of DOPS methods. Several participants provided positive comments about the feedbacks received from the clinical skills facilitators (CSFs). Results showed that DOPS assessment may be used as a successful tool in the assessment of preregistration house officers in hospital environment.

However, not providing the necessary trainings on how to take DOPS test, not being held at the scheduled time, not providing necessary feedbacks to participants and insufficient time for the test are the major weaknesses of the DOPS tests. Improving clinical skills and autonomy during the evaluation are the strengths of the method noted by participants in the reviewed articles. A study conducted by Naeem has also recognized DOPS tests as an effective tool to improve clinical skills (48). Some other studies have also mentioned this issue (31, 49-51). However, Bindal et al. study in the UK showed that DOPS tests cannot be used as a useful educational tool in improving practical skills (29). This could be due to the drawbacks of conducting DOPS tests, which were pointed out in the study by Bindal et al. According to them, the quality of conducting the tests was poor. Biased approaches towards participants and the stressfulness of the tests are the major weaknesses highlighted by previous studies (52-54). Thus, paying proper attention to the reliability of assessors, using selected multiple assessors for all participants, and videotaping can play a key role in addressing the problems involved in conducting this type of evaluation. Results of the studies conducted on the validity and reliability of DOPS tests have indicated an appropriate degree of validity and reliability. These findings are consistent with most studies undertaken in this field. Wilkinson et al. investigated the validity of DOPS tests from the perspective of the professionals and the result was good. Moorthy et al. also confirmed the validity of such tests in the field of surgery (55). Bould et al. reported that the validity of the tests in the field of anesthesiology is high (54). In addition, Weller et al. found that the reliability of these tests is high (56). Other studies have also reported an acceptable reliability (12, 57). Despite the good results of validity and reliability reported in studies, some resources have suggested conducting further studies on validity and reliability factors (58, 59). Because few studies have been conducted on the feasibility of this method, further studies are required in this area. Study results revealed that the satisfaction level of DOPS tests is fairly acceptable. Likewise, a study conducted by Shahid Hassan on surgery residents at the University of Malaysia indicated a satisfactory result with DOPS tests (60). Yoon Kang et al. at Weill Cornell Medical College suggested that the satisfaction level of students and professors was medium and high, respectively (25). The results of some reviewed studies showed that the satisfaction level of professors/assessors is higher than that of students/participants. This is consistent with the results of Harpreet Kapoor's study in India conducted on interns and professors of department of ophthalmology (61). One of the causes of dissatisfaction in conducting DOPS tests was insufficient time devoted to perform each skill. In many studies, this factor has been one of the main causes of student dissatisfaction (25, 55, and 62). Thus, it is recommended to review the procedure of the tests and devote more time to educational groups to perform each skill. Also, the results indicated that DOPS tests had a significant impact on improving student learning. The results of Holombae et al. study on medical students showed that those students who were evaluated by DOPS had a high skill level (63). Chen et al.

also suggested that DOPS tests in senior medical students have contributed to the increase of self-report, skill upgrading, as well as self-confidence (64). In a study in Taiwan, Tsui et al. stated that this type of test has a significant role in upgrading the skills and empowering medical students (35). As a result, it seems that in addition to being applied as a suitable method for evaluation purposes, DOPS can be used as an educational tool to educate and empower students. The study results revealed that the performance of participants who used DOPS was relatively satisfactory. Not providing the necessary trainings for taking DOPS test, not being held at the scheduled time, not providing necessary feedbacks to participants, and insufficient time for the test were the major drawbacks of DOPS tests.

One possible reason for the relatively good grades of participants on DOPS tests can be the motivation created by the tests. Another possible reason might be the easy approaches by assessors during the tests because the tests rely on instructor's evaluation method and their point of view. One of the limitations of this study was that the searches were limited to articles published in English and Persian languages. Also, many articles were excluded due to low quality.

Conclusion

Considering the high position of evaluation in medical education, an effective and appropriate method of evaluation is essential in this field. The results of this study revealed that DOPS tests can be used as an effective and efficient evaluation method to assess medical students because of their appropriate validity and reliability, positive impact on learning, high satisfaction level of students, and other advantages. However, special attention needs to be devoted to the quality of these tests.

Conflict of Interests

The authors declare that they have no competing interests.

References

1. American Education System – Education System in United States of America – USA Education System". Indobase.com. Retrieved 2010-04-14.
2. UNESCO, Education For All Monitoring Report 2008, Net Enrollment Rate in primary education.
3. UNICEF (United Nations Children's Fund). 2009. All children everywhere: A strategy for basic education and gender equality.
4. Special Education. Oxford: Elsevier Science and Technology. 2004.
5. Staff. Evaluation Purpose". designshop - lessons in effective teaching. Learning Technologies at Virginia Tech. 2011.
6. White R, Ewan CH. Clinical teaching in nursing. London: Chapman and hall; 1995.
7. Ghojzadeh M, Aghaei MH, Naghavi-Behzad M, Piri R, Hazrati H, Azami-Aghdas S. Using Concept Maps for Nursing Education in Iran: A Systematic Review. *Res Dev Med Educ*. 2014;3(1):67-72.
8. Vergis A, Hardy K. Principles of Assessment: A Primer for Medical Educators in the Clinical Years. *Inet. J. Med. Educ*. 2010;1(1).
9. U.S. Department of Health and Human Services. Physical Activity Evaluation Handbook. Atlanta, GA: Centers for Disease Control and Prevention, 2002. www.cdc.gov/nccdphp/dnpa/physical/handbook/pdf/handbook.pdf
10. Committee on Standards for Educational Evaluation. The Student Evaluation Standards: How to Improve Evaluations of Students. Newbury Park, CA: Corwin Press. 2003.
11. Kadagad P, Kotrashetti S. Portfolio: a comprehensive method of

- assessment for postgraduates in oral and maxillofacial surgery. *J Maxillofac Oral Surg.* 2013;12(1):80-4.
12. Al-Naami M. Reliability, validity and feasibility of the Objective structured clinical examination in assessing clinical skills of final year surgical clerkship. *Saudi Med J.* 2008;29(12):1802-7.
 13. Maylett T. 360-Degree Feedback Revisited: The transition from development to appraisal. *Compens. Benefits Rev.* 2009;41(5):52-9.
 14. Robinson B. The CIPP approach to evaluation, COLLIT Project. 2002.
 15. Jabbari H, Bakhshian F, Alizadeh M, Alikhah H, Naghavi Behzad M. Lecture-based versus problem-based learning methods in public health course for medical students. *RDME.* 2012;1:31-5.
 16. Ohn J, Norcini D. Assessment methods in medical education. *Teaching and Teacher Education.* 2007;23:239-50.
 17. Richard K, Reznick ME, MacRae H. Teaching surgical skills –changes in wind. *N Engl J Med.* 2006;355:2664-9.
 18. Moorthy K, Munz Y, Sarker S, Darzi A. Objective assessment of technical skills in surgery. *BMJ.* 2003;327(7422):1032-7.
 19. Sahebzamani M, Farahani H, Jahantigh M. [Validity and reliability of direct observation of procedural skills in evaluating the clinical skills of nursing students of Zahedan nursing and midwifery school]. *Zahedan. J. Res. Med. Sci.* 2012; 14(2): 76-81.
 20. Amin Z, Chong Y, Khoo H. Direct observation of procedural skills (chapter 16). In: Amin Z, Chong Y, Khoo H. *Practical guide to medical student.* 1st ed. Singapore: World Scientific Printers CO; 2006: 71-4.
 21. Wilkinson JR, Crossley JG, Wragg A, Mills P, Cowan G, Wade W. Implementing workplace-based assessment across the medical specialties in the United Kingdom. *Med Educ.* 2008; 42(4): 364-373.
 22. Norcini JJ, Mackinley DW. Assessment methods in medical education. *Teach Teacher Educ.* 2007; 23(3): 239-50.
 23. Asadi K, Mirbolook AR, Haghghi M, Sedighnejad A, Naderi nabi B, Abedi s, et al. Evaluation of Satisfaction Level of Orthopedic Interns from Direct Observation of procedural Skills Assessment (DOPS). *Res. Med. Educ.* 2012;4(2):17-23.
 24. Hoseini LB, Mazloum SR, Jafarnejad F, Foroughipour M. Comparison of midwifery students' satisfaction with direct observation of procedural skills and current methods in evaluation of procedural skills in Mashhad Nursing and Midwifery School. *Iran J Nurs Midwifery Res.* 2013;18(2):94-100.
 25. Kang Y, Bardes C, Gerber L, Storey-Johnson C. Pilot of Direct Observation of Clinical Skills (DOCS) in a Medicine Clerkship: Feasibility and Relationship to Clinical Performance Measures. *Med Educ Online.* 2009;14(9):1-8.
 26. Khoshrang H. Assistants' specialized perspectives on evaluation method of practical skills observations (DOPS) in Guilan University of Medical Sciences. *J. Res. Med. Educ.* 2011;2(2):40-4.
 27. Farajpour A, Amini M, Pishbin E, Arshadi H, Sanjarmasavi N, Yousefi J, et al. Teachers' and Students' Satisfaction with DOPS Examination in Islamic Azad University of Mashhad, a Study in Year 2012. *Iran J Med Edu.* 2014;14(2):165-73.
 28. Cohen SN, Farrant PB, Taibjee SM. Assessing the assessments: U.K. dermatology trainees' views of the workplace assessment tools. *Br. J. Dermatol.* 2009;161(1):34-9.
 29. Bindal N, Goodyear H, Bindal T, Wall D. DOPS assessment: a study to evaluate the experience and opinions of trainees and assessors. *Med Teach.* 2013;35(6):1230-4.
 30. Wilkinson J, Crolley J, Wragg A. Implementing workplace-based assessment across the medical specialties in the United Kingdom. *Med Educ.* 2008;42(4):364-73.
 31. Cobb K, Brown G, Jaarsma D, Hammond R. The educational impact of assessment: a comparison of DOPS and MCQs. *Med Teach.* 2013;35(11):1598-15607.
 32. Dabhadkar S, Wagh G, Panchanadikar T, Mehendale S, Saoji V. To evaluate Direct Observation of Procedural Skills in OBGY. *NJIRM.* 2014;5(3):92-7.
 33. Amini A, Shirzad F, Mohseni M, Sadeghpour A, Elmi A. Designing Direct Observation of Procedural Skills (DOPS) Test for Selective Skills of Orthopedic Residents and Evaluating Its Effects from Their Points of View. *Res Dev Med Educ.* 2015;4(2):1452-7.
 34. Akbari M, Mahavelati Shamsabadi R. Direct Observation of Procedural Skills (DOPS) in Restorative Dentistry: Advantages and Disadvantages in Student's Point of View. *Iran J Med Edu.* 2013;13(3):212-20.
 35. Tsui K, Liu C, Lui J, Lee S, Tan R, Chang P. Direct observation of procedural skills to improve validity of students' measurement of prostate volume in predicting treatment outcomes. *Urol Sci.* 2013; 24(3):84-8.
 36. Profanter C, Perathoner A. DOPS (Direct Observation of Procedural Skills) (in undergraduate skills-lab: Does it work? Analysis of skills-performance and curricular side effects. *GMS J. Med. Educ.* 2015;32(4).
 37. Scott DJ, Rege RV, Bergen PC, Guo WA, Laycock R, Tesfay ST, et al. Measuring operative performance after laparoscopic skills training: edited videotape versus direct observation. *J. laparoendosc. adv. surg. tech. a.* 2000;10(4):183-90.
 38. Shahgheibi S, Pooladi A, BahramRezaie M, Farhadifar F, Khatibi R. Evaluation of the Effects of Direct Observation of Procedural Skills (DOPS) on Clinical Externship Students' Learning Level in Obstetrics Ward of Kurdistan University of Medical Sciences. *J. Med. Educ.* 2009;13(1,2):29-33.
 39. Morris A, Hewitt J, Roberts CM. Practical experience of using directly observed procedures, mini clinical evaluation examinations, and peer observation in pre-registration house officer (FY1) trainees. *Postgrad. Med. J.* 2006;82(966):285-8.
 40. Hengameh H, Afsaneh R, Morteza K, Hosein M, Marjan SM, Abbas E. The Effect of Applying Direct Observation of Procedural Skills (DOPS) on Nursing Students' Clinical Skills: A Randomized Clinical Trial. *Glob. J. Health Sci.* 2015;7(7 Spec No):17-21.
 41. Nazari R, Hajhosseini F, Sharifnia H, Hojjati H. The effect of formative evaluation using "direct observation of procedural skills" (DOPS) method on the extent of learning practical skills among nursing students in the ICU. *IJNMR.* 2013;18(4):290-3.
 42. Bagheri M, Sadeghnezhad M, Sayyadee T, Hajiabadi F. The Effect of Direct Observation of Procedural Skills (DOPS) Evaluation Method on Learning Clinical Skills among Emergency Medicine Students. *Iran J Med Edu.* 2014;13(12): 1073-81.
 43. Watson MJ, Wong DM, Kluger R, Chuan A, Herrick MD, Ng I, et al. Psychometric evaluation of a direct observation of procedural skills assessment tool for ultrasound-guided regional anaesthesia. *Anaesthesia.* 2014;69(6):604-12.
 44. Barton JR, Corbett S, van der Vleuten CP. The validity and reliability of a Direct Observation of Procedural Skills assessment tool: assessing colonoscopic skills of senior endoscopists. *Gastrointestinal endoscopy.* 2012;75(3):591-7.
 45. Kuhpayehzade J, Hemmati A, Baradaran H, Mirhosseini F, Sarvieh M. Validity and Reliability of Direct Observation of Procedural Skills in Evaluating Clinical Skills of midwifery Students of Kashan Nursing and Midwifery School. *JSUMS.* ۰۵-۱۴۰۰(۱)۲۱;۲۰۱۴.
 46. Bazrafkan L, Shokrpour N, Torabi K. Comparison of the Assessment of Dental Students. *J. Med. Educ.* 2009;13(1,2):16-23.
 47. Mitchell C, Bhat S, Herbert A, Baker P. Workplace-based assessments of junior doctors: do scores predict training difficulties? *Med Educ.* 2011;45(12):1190-8.
 48. Naeem N. Validity, reliability, feasibility, acceptability and educational impact of direct observation of procedural skills (DOPS). *J Coll Physicians Surg Pak.* 2013;23(1):77-82.
 49. Nicol D, MacFarlane-Dick D. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Stud Higher Educ.* 2006;31:199-218.
 50. Swanwick T. *Understanding medical education: Evidence, theory and practice.* Wiley-Blackwell; Oxford: 2010.
 51. Norcini J, Burch V. Review Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach.* 2007;29(9):855-71.
 52. De Oliveria Filho G. The construction of learning curves for basic skills in anaesthetic procedures: An application for the cumulative sum method. *Anesth Analg.* 2002;95:411-6.
 53. Beard JD, Jolly BC, Newble DI, Thomas WEG, Donnelly J, Southgates LJ. Assessing the technical skills of surgical trainees. *Br J Surg.* 2005; 92(6): 778-82.
 54. Bould M, Crabtree N, Naik V. Assessment of procedural skills in anaesthesia. *Br J Anaesth.* 2009;103(4):472-83.
 55. Moorthy K, Munz Y, Sarker S, Darzi A. Objective assessment of technical skills in surgery. *BMJ.* 2003;327(7422):1032-7.
 56. Weller J, Jolly B, Misur M. Mini-clinical evaluation exercise in anaesthesia training. *Br J Anaesth.* 2009;102(5):633-41.
 57. Tudiver F, Rose D, Banks B, Pfortmiller D. Reliability and validity testing of an evidence-based medicine OSCE station. *Fam Med.* 2009;41(2):89-91.
 58. Review of work-based assessment methods. CIPHER. Australia: <http://mjiri.iums.ac.ir>

- USYD; 2007: 11-17.
59. Turnbull J, Gray J, MacFadyen J. Improving in-training evaluation programs. *J Gen Intern Med.* 1998;13:317-23.
 60. Hassan S. Faculty Development: DOPS as Workplace-Based Assessment. *J. Med. Educ.* 2011;3(1):32-43.
 61. Kapoor H, Tekian A, Mennin S. Structuring an internship programme for enhanced learning. *Med Educ.* 2010;44(5):501-2.
 62. Davies H, Archer J, Southgate L, Norcini J. Initial evaluation of the first year of the Foundation Assessment Programme. *Med Educ.* 2009;43:74-81.
 63. Holmboe E, Hawkins R, Huot S. Effects of training in direct observation of medical residents clinical competence: a randomized trial. *Ann Intern Med.* 2004;140(11):8781-4.
 64. Chen W, Liao S, Tsai C, Huang C, Lin C, Tsai C. Clinical Skills in Final-year Medical Students: The Relationship between Self-reported Confidence and Direct Observation by Faculty or Residents. *Ann Acad Med Singapore.* 2008;37(1):3-8.