

*Original Article***Evaluation of validity of midwifery special courses in Isfahan University of Medical Sciences***Ashraf Kazemi\**, *Soheila Ehsanpour\*\****Abstract**

**BACKGROUND:** The aim of training midwifery student is to increase scientific and practical abilities in trainees in order to present caring health services. Then, the evaluation tools of these abilities have great importance. This study tried to evaluate the content validity, criterion validity and base validity of academic exams.

**METHODS:** This cross-sectional research was an evaluated type that has been done on 18 special theoretical courses of midwifery in 2 semesters in 2007-2008. The data gathered by checklists. The data about questionnaire and the result of analyzing exam questions (final and midterm) were compiled by 2 educating experts of medical education and 2 experts for each course. The data analyzed by SPSS software. For determining the base validity, spearman correlation test and for presenting descriptive results, distribution tables were applied.

**RESULTS:** The evaluation of 1013 questions showed that in 18 courses, in 61.18% of exams, more than 90% of questions had content validity and in 28.27% of exams, in more than 90% of questions, criterion validity had been considered. The results showed that in 92.38% of questions, content validity and in 80.45% of questions, criterion validity was considered. 11 courses out of 18 courses had base validity.

**CONCLUSIONS:** This survey showed that content validity of the exam questions in midwifery special theoretical courses was in favorable levels. But, the criterion validity of exam questions was far away from the ideal level. Then, education in each session can help the teachers achieve their exam purposes.

**KEY WORDS:** Content validity, criterion validity, midwifery special exams, reliability.

IJNMR 2011; 16(2): 153-157

The profession of midwifery is the field that presenting its services with great bulk of knowledge and increasing the scientific abilities of the staffs has direct influence on mother and fetal health.<sup>1</sup> So, the training of midwifery students should be in a way to give services by using their theoretical knowledge and skills in their education periods.<sup>2</sup>

In every training plan, evaluation is essential that can promote the education from stationary to a dynamic mode.<sup>3</sup> In midwifery education, special theoretical courses and their evaluation have high importance for midwifery students in entering the clinical education.

Because the results of their evaluation determine the efficiency of students in giving midwifery services, it's needed to evaluate the efficiency of trainees. By evaluating the educational promotion of trainees and by comparing the recorded results with predefined educational purposes, it will be revealed weather the trainer and the trainees attempts have come to favorable results or not? So the quality of exam questions has a great importance. A good test is the one that enables us to evaluate and identify the efficiency base of educational contents.<sup>4</sup> Probably, evaluation is the most critical task of teachers; hence identifying the effective factors

\* Member of Teaching Staffs, Isfahan University of Medical Sciences and PhD Candidate in Reproductive Health, Tehran University of Medical Sciences.

\*\* MSc, Nursing and Midwifery Care Research Center, School of Nursing and Midwifery, Isfahan University of Medical Sciences, Isfahan, Iran.

Correspondence to: Ashraf Kazemi

Email: kazemi@nm.mui.ac.ir

Research Project of Isfahan University of Medical Sciences.

on efficiency and effectiveness of each question enables the teacher to design the exam with high rank of reliability.

Item analysis is a conglomeration of the methods by which the efficiency and the effectiveness of each question and the exam in whole are evaluated. Item analysis helps us study the features of each question and lead us to be insured that questions are in the level that can measure the expected abilities. It also lead us to apply vital tips in question structure which puts the answering as the sigh of learning and separates the students on the base of their efficiency and learning rate (the more active the student is, the higher scores he/she takes). For identifying this feature of the exam, the evaluation of base validity is applied by studying the correlation rate between students' scores with those of other courses.<sup>5</sup>

The importance of student's evaluation by the effective tools made the researches to evaluate the students measuring tools. Quality evaluating of multiple choice tests in medical school of Mazandaran University of Medical Sciences showed that out of 1478 questions in 25 exams, 64 questions had structural problems.<sup>6</sup> While Shakoornia showed that more than half of the questions designed by teaching staffs of Jondi-Shapor Medical University had correct structure with no fault.<sup>7</sup> Hadi's survey also showed that out of 111 medical trainees in Shiraz Medical School, no significant difference was reported between their knowledge and their scores in exams.<sup>8</sup>

All of these surveys indicated that the validity of exams questions in medical universities do not follow the same rate, and it is needed that every university evaluate the efficiency of students. The current condition of exams from different perspectives of evaluation has been studied based on content validity, structure validity and base validity in order to evaluate the efficiency of exams by post evaluation of collected information.

This paper tries to high lights the weakness spots in order to develop evaluation system of midwifery special courses.<sup>5</sup>

## Methods

This research was an evaluation survey. Cross-sectional study was carried out on 18 special theoretical courses of midwifery in 2 semesters in 2007-2008 in Isfahan University of Medical Sciences. The content validity, criterion validity and reliability of the exams were evaluated. The applied checklists were edited by the researchers of this survey and 2 educational experts in medical fields. The reliability of ideas of educational experts on evaluating criterion validity was evaluated by a pilot study. In this pilot study, the questions of final exams of 5 terms were distributed to 2 experts for evaluating the similarity of experts' evaluations. The results of evaluations (from different experts) were compared and confirmed with 0.72 alpha coefficients in correlation rate. The content validity of the questions on the base of question capability for measuring educational power of course plan were evaluated by the 2 experts in related course (not involved in exam designing). If a differentiation raised between the 2 experts ideas (in content validity), the question was evaluated by a third expert and just the similar results of the earlier experts were recorded. Criterion validity was evaluated on the base of Milman question designing principles.<sup>4</sup> Content validity and base validity were recorded as Valid or Not Valid.

The sampling type was census method. For data collecting, after holding the exams and releasing the students' scores to the education office of nursing and midwifery department, we referred to office staffs and clarified the purpose of this study and insured them about the secrecy of the raw data of each course including exams questions, course plan and students scores. The content validity and criterion validity of each question was evaluated by a course expert (teaching staffs of department) and a medical educational expert. Base validity of the courses was measured by the correlation of students' scores in considered course and corresponding course (obstetric course with gynecology for example) with the number of units and educational content. One of the features caused to

**Table 1.** Percent of questions with content and criterion validity in exams

Number of course	1	2	3	4	5	6	7	8	9
Content validity	100	100	78.3	80.5	83.3	97.5	100	95.1	94.6
Criterion validity	100	66.6	91.3	64.8	100	72.5	100	66.7	87.1
Number of course	10	11	12	13	14	15	16	17	18
Content validity	97.3	88	95	98	77.5	100	100	87.9	89.7
Criterion validity	94.8	54	78	72	80	89	92.1	17.5	51.7

withdraw a case from the survey was teachers' dissatisfaction in exposing the exam questions to us. The studied courses in this survey included: sexual disorder education and consulting, obstetric and related observations,<sup>1-4</sup> radiology and sinology in midwifery and gynecology physiopathology,<sup>1,3</sup> maternal and child health gynecology, genetics and psychology in midwifery, newborns, and pediatrics-midwifery management. The results of each course (by title) were not reported for loyalty considerations. The questions of analgesic and anesthesia were given to the researchers. Also, the exam questions of nutrition and embryology were not evaluated because we did not have access to their teachers. For data analyzing and presenting the results, prevalence distribution table and Spearman correlation were used.

## Results

The number of evaluated questions in 18 special theoretical courses in midwifery in 2 semesters was totally 1013. The average number of questions in each studied course was 54 (27.66). In 9 courses the questions were multiple choice types, in 3 courses the descriptive type was used, 5 courses included both multiple choice and descriptive ones and in 1 course multiple choice, descriptive and blank types were applied. The prevalence of questions with content validity and criterion validity is shown in Table 1.

The results in Table 1 show that in all of the exams, more than 75% of questions had content validity, in 61.11% of exams more than 90 of questions had content validity and in 28.78% of exams, more than 90% of questions had criterion validity. In 92.82% of total studied questions, content validity and in 80.45% of them criterion validity was seen.

For base validity determination, the correlation rate in students' scores in considered course and a corresponding course was evaluated by Spearman test. The results are shown in Table 2. The results indicate that 11 course exams had base validity.

## Discussion

The results that came from the current survey indicated that in all of the exams, approximately all of the questions had content validity; truly in 61.11% of exams, more than 90% content validity was shown. These results showed the consideration of question designers towards the truth that the questions should prepare the condition to measure the learning in an appropriate mode.<sup>5</sup> If a teacher wants to use the questions as an appropriate tool for guiding the student, it's necessarily to define a definite relation between teaching purposes and evaluation. Evaluation standard committee believes that the evaluating quality can be judged by its accuracy and appropriateness.<sup>9</sup> This survey showed that exam designers always take this point to the

**Table 2.** Correlation confidence of students' score from each course and equivalent course

Number of course	1	2	3	4	5	6	7	8	9
R	0.11	0.16	0.06	0.57	0.38	0.49	0.39	0.58	0.75
P	Ns	Ns	0.02	0.003	Ns	Ns	Ns	0.004	0.001
Number of course	10	11	12	13	14	15	16	17	18
R	0.89	0.83	0.75	0.49	0.58	0.31	0.65	0.38	0.75
P	0.001	0.001	0.001	0.02	0.001	Ns	0.002	Ns	0.001

Ns: no significant

considerations. But, the survey of Najari et al. showed that 45.7% of teachers in Ahvaz University of Medical Sciences evaluated the content validity of their questions.<sup>10</sup> Although in this survey, the validity of questions held an acceptable degree, the criterion validity of them were not as favorable as content validity. Actually in 28.78% of exams, more than 90% criterion validity was seen. Vakili et al<sup>11</sup> in Kashan University of Medical Sciences and Tabatbaei et al<sup>12</sup> in Mashhad University of Medical Sciences showed the same results. The lack of structural principles in question designing disrupt structural validity and this lead the students to answer the questions not by scientific knowledge but by guessing methods. This will disrupt the students' attitudes towards exam. Toghyanifar et al survey which dealt with this problem, showed that the disruption of item analyzing principles in considerable number of multiple choice exams had a negative effect on test quality.<sup>13</sup> McCoubrie<sup>14</sup> in Bristol University and Hammond et al<sup>15</sup> in Hampton University reported problems in considerable number of questions.

However, the results of Shakoornia in Ahvaz revealed deferent results.<sup>7</sup> But, according to the current evidences in this survey and other studies, the empowerment and updating standardized plans and promoting the evaluating quality of trainees is essential for evaluation. It's also necessary to develop teachers' knowledge in the field of student evaluation, in designing and holding the exams processes.

The results of base validity measurement indicated that tests of 61.11% of courses had base validity. The prevalence of 38.89% in evaluations without base validity showed the

weakness of evaluation system by base validity perspective in special theoretical courses of midwifery. Also, Najari reported that just 10 teachers reviewed base validity and 7.5% rechecked criterion validity of questions, while the knowledge of 41.15% of teachers about base validity and 42.8% of them about criterion validity were enough.<sup>10</sup> The high rate of prevalence of questions with no base validity indicated the fault in designing multiple choice questions; it also showed that question designers of these tests had no good consideration in suitable answers, as the good students choose those. So it's needed that the question designers remove these faults by evaluating base validity of their questions. The results of the paper let us know that criterion and base validity of the questions suffered from some weaknesses. It also informed us that for item analyzing and its promotion, a correlating link should be existed between technical structure of education (like medical education development center) and the teachers.

It also advised that the designers analyze the questions and by using the results, recommends about the quarries and take good actions for promoting the questions quality for further usage. Clearly, by stabilizing this method, the knowledge and the reliability of teachers in designing appropriate tests adopted with standard principles will be developed. This is one of the most reliable methods in determining, developing and promoting education evaluating quality.

The authors declare no conflict of interest in this study.

## References

1. Smith-Stoner M. Critical thinking activities for nursing. 4th ed. ST Louis: Mosby; 1999.
2. Klass D. Reevaluation of clinical competency Am J Phys Med Rehabil 2000; 79(5): 481-6.
3. Wink DM. Using questioning as a teaching strategy Nurse Educ 1993; 18(5): 11-5.
4. Seyf AA. Educational measurement, assessment and evaluation. 4th ed. Tehran: Doran, 2003.
5. Naderi E, Seyf Naraghi M. Evaluation and measurement and basic analysis tools in cultural sciences and psychology. 1st ed. Tehran: Misagh; 1997.
6. Haghshenas MR, Vahidshahi K, Mahmoudi M, Shahbaznezhad L, Parvinnezhad N, Emadi A. Evaluation of Multiple Choice Questions in the School of Medicine, Mazandaran University of Medical Sciences, The First Semester of 2007. Strides in Development of Medical Education 2008; 5(2): 120-7.

7. Shakoornia A, Khosravi A, Shariati A, Zarei A, editors. Survey on multiple choice questions of faculty members of Jondi-Shapor Medical University of Ahwaz. Proceedings of The 8th National Congress of Medical Education; 2007; Kerman, Iran: Kerman University of Medical Sciences; 2007.
8. Hadi N, Shiraly R. Medial Interns' knowledge about Hypertension in Shiraz University of Medical Sciences. Iranian Journal of Medical Education 2005; 5(2).
9. Sabsevari S. Meta-evaluation and its Standards. Strides in Development of Medical Education 2005; 2(2).
10. Najari SH, Abedi P. Knowledge and practice of faculty members in assessing the validity and reliability of examinations. The Journal of Qazvin University of Medical Sciences 2004; 30(7).
11. Vakili Z, Sayah M, Fakhrian E, Rasouli Nejad SA, Mosayebi Z, Moniri R. editors. Survey on psychometric characteristics of resident examination Kashan Medical University. Proceedings of The 8th National Congress of Medical Education; 2007; Kerman, Iran: Kerman University of Medical Sciences; 2007.
12. Tabatbaei K, Bahreyni Tousi SMH, Derakhshan A, Khajeh Dalouei M, Gholami H. Analytical assessment of multiple-choice tests/at Mashhad Medical School. Medical Journal of Mashhad University Of Medical Sciences 2002; 45(76): 89-95.
13. Toghyanifar N, Gorgyaraghi M, Esmaili A, Avizhgan M, Mazaheri M. editors. Survey on quality of Multiple Choice Questions of clinical clerkship Isfahan Medical Faculty 2005–2006. Proceedings of The 8th National Congress of Medical Education; 2007; Kerman, Iran: Kerman University of Medical Sciences; 2007.
14. McCoubrie P. Improving the fairness of multiple-choice questions: a literature review Med Teach 2004; 26(8): 709-12.
15. Hammond EJ, McIndoe AK, Sansome AJ, Spargo PM. Multiple-choice examinations: adopting an evidence-based approach to exam technique Anaesthesia 1998; 53(11): 1105-8.