[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

Journal of Medical Education Development, 2017, Vol 10, No 27, 22 - 37
http://zums.ac.ir/edujournal/

Zanjan University of Medical Sciences

Comparison the Efficiency and the Effectiveness of a New Comprehensive Test Management Software (NAJMA Software) with Traditional Method in Resident in-training and pre-board Examinations in Isfahan University of Medical Sciences

Shahram Shayan[1], Masih Saboori[2], Omid Aghadavoudi[3*], Masoud Borumand[4], Habibollah Rezaei[5]

[1]Medical Education Department, Isfahan University of Medical Sciences, Isfahan, Iran.
[2]Neurosurgery Department, Isfahan University of Medical Sciences, Isfahan, Iran.
[3]Anesthesiology Department, Isfahan University of Medical Sciences, Isfahan, Iran.
[4] Faculty of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran.
[5]Medical Education Department, Isfahan University of Medical Sciences, Isfahan, Iran.

Article Info

Article Type:
Original Article

Article history:
Received 1 Feb 2017
Accepted 15 Aug 2017
Published17 Mar 2018

Keywords:
Efficiency
Effectiveness
Resident Evaluation
Comprehensive Test Management Software (CTMS)
NAJMA Software

Abstract

Background and Object: Considering new e-teaching and e-learning systems, and electronic assessment tools, this study was performed assessing the efficiency and the effectiveness of a new generated comprehensive test management software (CTMS), named NAJMA, in medical resident evaluation.

Materials and Methods: This cross sectional descriptive study was conducted in Isfahan University of Medical Sciences in 2015. Statistical analysis of the questions was done comparing the two models of examination management performed in 2014 and 2015 to evaluate all medical residents of 21 specialties in Faculty of Medicine. Totally 6300 questions from 42 question books based on the model of exams hold in 2014 and 2015 were analyzed. A survey questionnaire was administered to faculty members to obtain their insight into their perceptions of the two models.

Results: The cost of formatting and performing the examinations was considerably lower in CTMS method. From the whole 377 members. 64.8% of faculties preferred the CTMS method in management of resident examinations. Number of residents participated in the examinations was 711 in 2014 and 668 in 2015. There were no differences among traditional and CTMS methods considering statistical analysis and Millman structural criteria.

Conclusion: Improving and utilization of CTMS method in all university and state examinations will result in reducing human and material costs. Meanwhile satisfaction of faculty members and security of the examinations will be improved by eliminating intermediate processes from developing to analyzing the examination questions.

[*]***Corresponding Author***: Omid Aghadavoudi, ***E-mail***: aghadavoudi@med.mui.ac.ir

This article is referenced as follows: Shayan S, Saboori M, Aghadavoudi O, Boroumand M, Rezaei H. Comparison the Efficiency and the Effectiveness of a New Comprehensive Test Management Software (NAJMA Software) with Traditional Method in Resident in-training and pre-board Examinations in Isfahan University of Medical Sciences. J Med Educ Dev. 2018; 10 (27) :19-31

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

Introduction

Nowadays, there are many advances in medical education, and new electronic methods of education and learning have emerged. Therefore, in the field of assessment, it is required to use computer and electronic methods in designing and executing exams, and try to develop the designed software and improve the quality and efficiency of the system through evaluating them. Now, multiple choice questions examinations are widely used in mid-term and final assessment of different fields such as medical resident evaluation exams. Studies indicate the necessity of more active interaction of technical-educational structures of universities such as education development and assessment committees of the faculty with faculty members in regards to fundamental measures before and after exams including precise analysis of exams and investigating the accuracy and efficiency of questions.

On the other hand, in higher education systems of the state, assessment and examination is a serious challenge in front of the medical education authorities, and one of the main demands of the community is paying attention to these exams and their security and accuracy. Efficiency is a concept that evaluates the cost of resources spent on purpose achievement process. Effectiveness is the level of achievement to the determined purposes. In other words, effectiveness indicates that how much of the attempts have had the considered results. In terms of cost-effectiveness, some researchers have compared traditional methods such as paper based methods with electronic methods, and different results have been achieved. Analysis of multi choice questions is among categories which suggest the necessity of using software. In a study, using quantitative analysis of questions has been considered as a factor of improvement of reliability coefficient of exam. In a study conducted by Kazemi, in a significant portion of exams, the level of designed questions did not match the level of expected learning. In studies conducted by McKorbi et al. in Bristle University, and Hammond et al. in Hampoon University, a significant percentage of questions reported the existence of some problems. Shakkornia et al. investigated assistance exam of Jondi Shapoor medical sciences university and concluded that these questions have structural problems, which most of them have been prepared in low cognitive levels, so they have declared that the questions of these exams needed to be revised and changed. The results

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

of the studies conducted by Big et al. have indicated that multi choice questions can assess higher levels of learning compared to short answer questions, however, in this kind of questions, there are more faults in the choices.

In this regard and in order to solve many problems, a comprehensive test management software (CTMS), named NAJMA has been created which is designed based on the experiences of Iranian assessors with the purpose of current concerns of the country. This software has been designed based on Word software environment, and the test maker logs in with his/her username and private password, and after typing the question and choices, saves questions in the computer while removing structural faults in terms of Millman principles. Then, the test maker sends the coded questions to the central server via university VPN. Among other features of this software, the possibility of budgeting the questions according to the purposes and plans of the course can be pointed. The designers of this software believe that this software is able to perform all the steps of exam including designing questions, creating question bank, choosing exam questions and creating books with different arrangements and executive phases of tests such as printing, copying, correcting and scoring the answer sheets and final phases including analyzing the results and evaluating questions and reports.

In traditional methods, the test maker creates multi choice questions and the test maker or the secretary types the questions and finally the questions are printed in examination books. In this method, there is a smaller chance of preparing question banks and using them and also control accurate criteria of question designing. Whereas in electronic methods, the software is able to control the defined standards for questions and prepare question banks and extract different analyses and categorizations on questions. Therefore, the purpose of this study is to investigate the effectiveness and efficiency of using comprehensive test management software (CTMS) in comparison with the traditional methods in resident in-training and pre-board examinations in Isfahan School of Medicine.

Materials and Methods

This research is a cross sectional descriptive study which was conducted in 2015. In this study, efficiency and effectiveness of resident in-training and pre-board examinations in School of Medicine of Isfahan University of Medical Sciences were

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

investigated through two methods of traditional in 2014 and CTMS in 2015. Evaluation of efficiency was performed according to the assessment of costs and comments of faculty members, and evaluation of effectiveness was performed according to the assessment of examination quality indexes and also quantitative analysis of the exam.

Data collection was performed in three parts: A) Costs of the two methods of traditional method in 2014 and CTMS method in 2015 were evaluated by the documents available in financial department of education deputy of the university and specialist deputy of school of medicine. B) Comments of faculty members of clinical groups in regards to traditional and CTMS methods were collected by a questionnaire containing 16 main items which were approved by medical education and clinical group faculty members of school of medicine in terms of validity and reliability. The questionnaire was prepared in form of a qualitative poll checklist regarding the comparison between traditional and CTMS methods in terms of quality of exam questions, time, security and convenience of preparation, edition and finalizing the questions, and the answers were aggregated in the form of relative frequency (table 2). Content validity of the questionnaire was approved by 5 medical education specialists. In order to assess the internal coordination of questions, that is validity of the questionnaire of the faculty members, due to having three-choice questions, Cronbach's alpha method was used in the performed pilot on 10 peoples of the studied population and the result was 0.83 which shows the desirable reliability of this tool. The time required for filling the questionnaire was 10 minutes which was performed at the same time if the faculty members were present and willing to do so. In the case of the faculty members' absence or needing time, the questionnaire was delivered to the faculty members by the secretary of department and token back in the due time. At the end of the questionnaire, the faculty members were asked a question as the following: if you have any explanation, recommendation or criticism about the method of the exams, please write it. In several cases, faculty members stated descriptive explanations which are investigated in discussion section. C) Qualitative indexes of resident evaluation exam were obtained from the documents of analysis of assistance examination results in 2014 and 2015. Also, cumulative checklist of the quantitative and qualitative results analysis for the 21 clinical groups of school of

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

medicine was designed by the researcher and used by the faculty members of medical education group. Qualitative analysis of the questions was performed by the researcher and according to Millman criteria and study of the resident in-training and pre-board examinations questions in 2014 and 2015. About 6300 questions of 42 books that was 3150 questions of 21 books in 2014 were analyzed through traditional method, and 3150 questions of 21 books in 2015 were analyzed through CTMS method.

The 21 clinical groups of school of medicine included cardiology, orthopedics, emergency medicine, pediatrics, internal medicine, neurology, obstetrics and gynecology, pathology, ophthalmology, infection medicine, urology, physical medicine and rehabilitation, social medicine, neurosurgery, ear- nose and throat surgery, psychiatry, dermatology, anesthesiology, radiology, general surgery, and radiotherapy. The criteria for joining the study for the faculty members included their interest to engage in research and having two years of experience in working with traditional and CTMS software methods. In the phase of evaluating qualitative indexes of exam, groups with less than 20 assistants were omitted due to unsupported results (non extensibility in terms of statistics).

Research variables included three main categories of calculated costs, comments of faculty members (about superiority of traditional or CTMS methods in every questioned area and demographic particulars of faculty members) and qualitative indexes of exam (like difficulty index, differentiation index and observance of Millman criteria in designing questions).

Data analysis was performed using version 22 of SPSS software. Descriptive statistics were presented in form of mean, standard deviation and median. In order to test the hypotheses, suitable statistical tests including independent t were used in quantitative variables such as difficulty coefficient and differentiation coefficient of questions and Chi-square in qualitative and rank variables. Significance value was considered as $P< 0.05$.

Results

The results of evaluation of efficiency were obtained according to the assessment of costs and comments of faculty members as the following:

A) Costs: costs of resident in-training and pre-board examinations in traditional method

in 2014 and CTMS method in 2015 is presented in table (1).

Table 1: Costs of resident in-training and pre-board examinations in traditional method in 2014 and CTMS method in 2015

Costs	**Traditional**	**CTMS**
Faculty members accommodation	81725584	29847000
Faculty members food	170600000	43405000
Faculty members transportation	18067500	5724000
Personnel's salary	241150000	185000000
Total costs	511543084	263976000

All numbers are presented as Rials.
CTMS: comprehensive test management software

B) Facultymembers' poll: in faculty members' poll and studying their points of view in regards the two examination methods, from 377 members of clinical groups of school of medicine as the research community, 125 questionnaires were collected which showed a 33% participation of faculty members in answering the questionnaires. Among them, 75/6% were men and 24/4% were women. Among the participants, 14/6% were proffesors, 43/1% were associate professors and 42/3% were assistant professors. The average age of the faculty members was 47.6 ± 6.4, with the minimum of 38 and maximum of 62.

Working background of the faculty members was 13.4 ± 7.6 years with the minimum of 2 and maximum of 35. In regards to the clinical group faculty members' points of view and level of satisfaction about the method of examination, the answers of the faculty members of different groups about the comparison of traditional and CTMS methods in examination are presented in table 2. In 2014, the faculty members spent an average time of 2.6 ±2.9 hours for each question in traditional method (0.5 to 20 hours) and 3.4 ± 4.4 hours in 2015 (0.5 to 30 hours).

The results of evaluation of effectiveness were obtained according to assessment of qualitative indexes of exam and qualitative analysis of exam questions as the following:

C) Qualitative and quantitative indexes of exam questions: the number of participants of the test was 711 residents in 2014 and 668 residents in 2015. The results of the cumulating checklist, the results of the quantitative analysis for the 21 clinical groups of the school of medicine are presented in table 3. According to Kuder–Richardson Formula 20 (KR-20), reliability coefficient of the exam was 0.85±0.05 on average in the

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]
[DOR: 20.1001.1.22519521.1396.10.27.6.2]
[DOI: 10.29252/edcj.10.27.19]

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

traditional method and 0.86±0.05 in DTMS method, that according to t test, there was no difference between the two groups (P>0.05). However, in groups having less than 30 and particularly less than 20 participants, the final interpretation had to be done carefully, so groups with less than 20 people were omitted from the calculations. The percentage of correct answers to exam questions (Mean P) was 0.61 ± 0.05 in the traditional method and 0.62 ± 0.05 in CTMS method, so they did not have a significant statistical difference. Differentiation index of exam questions was 0.21 ± 0.03 in the traditional method and 0.21 ± 0.04 in CTMS method, without any significant difference. The number of suitable questions was on average 29 ±7 in the traditional method and 31±9 in CTMS method, without any statistically significant difference. Suitable questions were questions with a P value of between 0.3 and 0.7 and a differentiation index of more than 0.3.

Table 2: Comparison of faculty member' points of view and level of satisfaction about traditional and CTMS methods

	Item	Traditional	CTMS	No opinion
1	Quality of questions	47.2	51.2	1.6
2	Security of preparing, editing and finalizing questions	26.4	72	1.6
3	Speed of preparing, editing and finalizing questions	56	44	0
4	Convenience of preparing, editing and finalizing questions	60.8	36.8	2.4
5	Access to questions bank of previous years	19.2	78.4	2.4
6	Feasibility of editing questions based on principles of Millman	23.2	72.8	4
7	Attention to budgeting the questions according to blue print	23.2	72.8	4
8	Costs of preparing questions and exam	37.6	51.2	11.2
9	Feasibility of preparing question bank	16.8	77.6	5.6
10	Feasibility of searching questions by key words	16.8	83.2	0
11	Better multitask management for faculty members	42.4	57.6	0
12	Reducing the number of objected questions	35.2	48.8	16
13	Reducing the number of deleted questions	32	54.4	13.6
14	Reducing the typewriting errors in questions	17.6	69.6	12.8
15	Do you agree with which method for future exams?	35.2	64.8	0

All numbers are presented as percent. Number of faculty members who participated in completing the questionnaire were 125.

CTMS: comprehensive test management software

Table 3: Results of the cumulating checklist, the results of the quantitative analysis for the 21 clinical groups of the school of medicine

Item	Traditional	CTMS	t- value	P- value
Mean number of residents participating in the exam in each clinical group	34±21	32±18	1.12	0.276
KR20 reliability	0.85±0.05	0.86±0.05	0.267	0.792
Standard error	4.9±0.2	4.9±0.2	0.113	0.911
Mean P	0.61±0.05	0.62±0.05	0.725	0.477
Mean RPBis	0.21±0.03	0.21±0.04	0.235	0.817
Number of suitable questions	29±7	31±9	0.117	0.908

KR20 reliability: reliability of exam, Mean P: percent of correct answer to questions, Mean RPBis: differentiation index of exam

CTMS: comprehensive test management software

In qualitative analysis, every examination books containing 150 four-choice questions of 21 clinical fields were evaluated by a traditional method in 2014 and CTMS method in 2015 in terms of observance of structural principles of Millman. In other words, about 3150 questions in every method and about 6300 questions, totally, were investigated. Study of the observance of structural principles of Millman was performed based on the adjusted list of the executers as the following. The list included incomplete stem of the question, using vague terms and symbols in stem, multiplicty of evaluated educational purposes, association of the question stem with key choice, not underlining the negative verb in the question, using double negative clauses in the question and choices, heterogeneous choices with respect to length, grammar or language, using unnecessary and vague clauses in question, using all or non-clauses in choices, repetitive words in question and choices, and contradictory choices. In general, there was not any difference between the two methods with respect to the observance of structural principles of Millman. A total number of structural faults in questions was 44 in the traditional method and 39 in CTMS method. In addition, the number of typewriting errors in examination books of the 21 fields was studied in the two methods. This number was nine in the traditional method and 8 in CTMS which were almost equal.

The results of investigating residents' objections to the exam of 2014 and 2015

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

respectively through methods of traditional and CTMS are presented in table 4. The average number of objected questions in each field of study was 16.6±11.9 in the traditional method and 14.4±7.9 in CTMS method, without any statistically significant difference. The average number of omitted questions in each clinical group was 2±1.8 in traditional and 2.3±2.5 in CTMS method, without any significant difference. The average number of questions having two answers after correction of exam key were 2.4±1.8 in the traditional and 3.3±3 in CTMS method. The average number of questions in which the key had been changed was 1±1 in traditional and 1±0.7 in CTMS method. In general, the average number of accepted objections in each clinical group was 5.5±2.8 in the traditional method and 6±3.7 in CTMS method. In all mentioned items, there was not any statistically significant difference between traditional and CTMS methods.

Table 4: The results of investigating residents' objections to the exams with traditional and CTMS methods held in 2014 and 2015 respectively.

Item	Traditional	CTMS	t- value	P- value*
Objected questions	350±11.9	304±7.9	0.739	0.468
Omitted questions	43±1.8	50±2.5	0.681	0.5.4
Questions having two answers	51±1.8	70±3	1.28	0.213
questions with changed key answer	23±1	16±1	1.099	0.285
Total accepted objections	117±2.8	126±3.7	0.551	0.588

Data are presented as number of questions out of 3150 questions in each method.

*According to t- test

CTMS: comprehensive test management software

Discussion

In this study, efficiency and effectiveness of the traditional and CTMS methods in 2014 and 2015 for resident in-training and pre-board examinations were investigated in the school of medicine of Isfahan University of Medical Sciences. Efficiency was evaluated based on an assessment of costs and also comments of faculty members and effectiveness were evaluated based on the assessment of test quality indexes and quantitativeanalysis of examination questions. A) Costs: the results of the present research showed that costs of examination in traditional method was twice more than CTMS method. However, regarding the annual inflation and comparison between the two years, the observed difference has some

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

specific remarks. These differences are observed higher in sections in which there is a need for physical presence and typing by the typist. One of the advantages of electronic exams is that intermediate systems and people are omitted, and in addition to less probability of a fault in the process of information transfer between the phases, time and costs of intermediate phases are saved. Also, with reduction of the number of individuals involved in the exam, the privacy condition of the examination is increased. However, in CTMS method some hidden costs such as network costs, software, and hardware costs and some of these costs are higher in the first execution, and in the case of the popularity of them in universities, would be decreased. In the research executed by Mandel, the costs of execution of exams including designing and correcting questions in paper based methods required more time and expenses compared to computer methods. In contrast, in the research executed by Lim et.al in Singapore, using a computer in explanatory exams required more time and people than paper based methods. However, the researcher preferred this method due to its various advantages and more satisfaction of students and personnel. Side costs which are sometimes less considered such as the costs of faculty members and personnel's food and transportation have constituted a major part of expenses in the investigations. In executing exams by software methods, much of these costs are decreased. In this research, in faculty members' questionnaires, 51% of them have considered exam execution by CTMS more economically, and only 37/6% have considered the traditional method less expensive. However, it should be accepted that in CTMS method, a primary cost is required to prepare CTMS software, but as this software can be used for all subsequent examinations, the cost of preparing the software will be divided, but network, software and hardware costs will remain. In addition to above mentioned tangible costs, other intangible and hidden costs such as reduction of time spent for finalizing questions by the faculty members, and also decrease of the time spent by the exam executors and designers and execution of exams are among the merits of CTMS method. Also, removing paper costs of repeated check prints in traditional method leads to lower costs and protection of the environment by reducing cutting trees.

B) Faculty members' polls: in faculty members' polls and assessment of their satisfaction and attitudes towards the two

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

methods of exam execution, half of the faculty members regarded the exam question of a higher quality in the traditional method. Also, in regards to speed and convenience of preparing, editing and finalizing the questions, more than half of the faculty members preferred the traditional method. This can be attributed to the novelty of CTMS software because every new change is faced with difficulties in being accepted and institutionalized. In addition, some faculty members have limited skills in typing rather than writing on paper. So, they have spent more time on every question in CTMS method. Also, it seems that in the traditional method, the faculty members are less dealing with other phases of edition and finalization of exam questions, and have less observed the problemsand difficulties of them.

In regards to the possibility of access to designed questions of previous years, most faculty members have preferred CTMS method. Also, most of them have preferred the possibility of editing questions based on standard criteria of Millman and budgeting of questions according to the blue prints in CTMS method. Faculty members of the school of medicine have preferred CTMS method rather than traditional method in cases such as the possibility of creating question bank and searching the question of a certain topic using keywords (table 2). In the researcher's idea, the possibility of gaining access to the question banks of previous exams with the practical approach applied in CTMS method can lead to the idea of designing new questions by the test makers and increasing the reliability of exams by entering questions with high differentiation power in the process of question designing.

In the open question at the end of the questionnaire, the faculty members defending traditional method pointed to the following faults regarding CTMS method. Some of them regarded the questions designed in CTMS of less quality because they believe that since the faculty members should type the questions in CTMS by themselves, use short questions with short choices which leads to lower quality of questions. Also, they believe that regarding the experimental schedule of Isfahan University of Medical Sciences for finalizing question by a three member group, they have less access to the faculty members making questions at the time of reviewing questions. Also, some of the faculty members believe that editing of questions is more difficult in CTMS software and some consider the CTMS time consuming. Some have stated that DTMS has software problems

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

and working with this software is not easy and there is a need to revision by specialist persons. The faculty members defending CTMS method in the execution of residents' exam, declared that in order to use and institutionalize CTMS in all levels for important university exams, there should be wide efforts. Some of them believe that the most important problem is concerned with lack of skill in typing and entering the questions in the software, and through gaining experience and skill, it is possible to promote this method.

However, in this research, 64.8% of the faculty members agreed with execution of exams in CTMS in the subsequent upgrade and certification exams, and only 35.2% suggested the traditional method. Particularly, most of the faculty members prefer the CTMS method for its possibility of time and coordination with other faculty responsibilities. In the research executed by Maru et.al, a questionnaire concerning the satisfaction of students and faculty members about using this software in the evaluation of the dentistry faculty students was prepared and sent for all of them by e-mail. The level of participation of the faculty members and students were respectively 79% and 56%. According to the answers, it was evident that most of the students and faculty members preferred electronic evaluation and scoring rather than paper based methods. By referring to the available documents and interviewing the people involved in CTMS software development, the researcher has found that the main committee of designing and running the software has recognized major problems of the software through getting feedbacks from test makers and running CTMS software in important national exams such as assistants and certification exams.

C) Qualitative and quantitative indexes of the exam questions: in cumulative checklist of quantitative analysis results, the mean reliability coefficient of exams (KR 20 reliability) in traditional and DTMS methods is desirable and above 0.8. However, since the participants of resident evaluation exam are from different levels, for example first, second and third and fourth years, the reliability coefficient of the exam is increased. In addition, high number of questions in residents' exams (about 150 questions) is another cause of increase of reliability coefficient. The mean percentage of correct answers to the exam questions (Mean P) in traditional and CTMS methods are equal and about 0.6. Generally, suitable P values in academic achievement exams are between 0.4

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

and 0.8. The mean differentiation index of exams in traditional and CTMS methods are similar and about 0.21. The mean number of suitable questions in the exam in traditional and CTMS methods are similar.

In evaluation of cumulative checklist of qualitative analysis results for the 21 clinical groups in accordance with Millman's structural principles, it was expected that CTMS method is better than traditional method having the capacity of automatically checking Millman's structural principles. However, in general there was no difference between the two methods in terms of observance of Millman's structural principles. These results show that the faculty members who observed Millman's structural principles in traditional method, have a knowledge of principles of multiple choice question designing, and in the case of any problem, the questions have been improved by the examination board inspection groups. On the other hand, regarding the approach of Isfahan University of Medical Sciences considering the presence of a scientific committee composed of specialist test makers at the time of designing and finalizing the questions and instantaneous feedbacks at the time of finalizing the questions from the beginning of the process of delivering these exams to the universities, the structural problems have been minimized. In several internal and foreign studies on the quality of multi-choice questions of the exams in schools of medicine and paramedicine in previous years, the structural problems of these questions have been reported about 33 to 46 percent. However, in this research, the relative frequency of questions having structural problems in traditional and CTMS methods are respectively 1.39 and 1.23 which is desirable. This difference might be due to the supervision of scientific committee in Isfahan University of Medical Sciences. In addition, the number of typewriting errors in examination books of the 21 fields was investigated in the two methods and this number was 9 in traditional method and 8 in CTMS method that were almost equal. The significant point is that in the traditional method, a skillful typist types the questions, whereas in CTMS method, the faculty member types the questions. Although in both methods the number of typewriting errors is observed very few in regards to 6300 questions, and this is desirable. Also in traditional method, typewriting errors were corrected after repeated check prints which require more time and costs, whereas in CTMS method, this is performed at the time

of designing questions or reviewing by a three member group. However, it is noticeable that in faculty members' poll questionnaire, 69.6% of the faculty members believed that the number of typewriting errors was fewer in CTMS method. In evaluation of the assistants' objection to the traditional exam of 2014 and CTMS exam of 2015, the mean number of stated and accepted objections did not have any relationship with the exam method. One of the restrictions of this research was that it was not possible to access the blue print tables of 21 fields of school of medicine and comparing the taxonomy of the designed questions with blue print tables, due to their privacy. Therefore, it is suggested that in evaluation of the questions, the taxonomy of designed questions be adapted to the blue print of every clinical group.

The other limitation of this research is the low percentage of (33%) the faculty members' participation in answering the questionnaire and the reasons need to be reflected. Beside these restrictions, there are some strength points in this research: this research is the first comparative investigation which evaluates the efficiency and effectiveness of CTMS software after running in universities and compares it with traditional method of designing, execution and analyzing examinations. In qualitative and quantitative analysis of the questions, all the questions of examination books of the 21 clinical groups have been investigated and the two methods were compared in a very precise manner. It is suggested that in execution of all academic achievement tests and national exams such as resident acceptance and certificate exams, the CTMS method be used instead of traditional methods, due to its efficiency and reliability.

In addition, after execution of several exams by CTMS software, in future investigations, faculty members' polls and qualitative and quantitative analysis of the questions are done, so that after improving software problems and more familiarity of the faculty members with this method, a more realistic assessment of the new condition is performed. Besides, in the case of need to use the facilities of sending questions in virtual private network (VPN), this requires a high security of network communications. The advantages of this software include the possibility of preparation of question bank, adapting the questions with blueprint, and searching the questions of a certain topic according to keywords. In addition, it is possible to prepare several examination books with similar contents and different question sequence, which increases the security

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]
[DOR: 20.1001.1.22519521.1396.10.27.6.2]
[DOI: 10.29252/edcj.10.27.19]

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

coefficient of the exam. Since creating good questions banks requires the cooperation of all faculty members of a department and even several similar departments of different universities, using this software, we created question banks across the country, and this has being performed in 2016 and 2017. Also, the possibility of using exam resources as electronic books for quick access to test makers has been provided. However, electronic methods such as CTMS require gaining experience and skill in designing and running by faculty members in workshops.

Conclusion

Execution of academic exams, such as resident in-training and pre-board examinations using CTMS software can have a higher efficiency than traditional methods because it can lead to decreasing costs and time spent by the faculty members and exam executors in long term. Furthermore, the effectiveness of the new method is higher than traditional method, because there is more possibility of satisfaction of clinical departments' faculty members in this method.

Acknowledgments

This research was financially supported by Isfahan University of Medical Sciences. Hereby, we extend our gratitude to all of the faculty members and education authorities and Computer Department of School of Medicine in Isfahan University of Medical Sciences and the faculty members of Medical Education Department for their cooperation with the research.

References

1- Arevalo CR, Bayne SC, Beeley JA, et al. Framework for e-learning assessment in dental education: a global model for the future. *J Dent Educ* 2013; 77(5):564-75.

2- Karay Y, Schauber SK, Stosch C, Schuettpelz-Brauns K. Can computer-based assessment enhance the acceptance of formative multiple choice exams? A utility analysis. *Med Teach*. 2012; 34(4):292-6.

3- Mc Courbie P. Improving the fairness of multiple choice questions: A literature review. *Med Teach*. 2004; 26(8): 709 –712.

4- Mandel A, Hörnlein A, Ifland M, Lüneburg E, Deckert J, Puppe F. Cost analysis for computer supported multiple-choice paper examinations. *GMS Z Med Ausbild*. 2011; 28(4):Doc55.

5- Lim EC, Seet RC, Oh VM, Chia BL, Aw M, Quak SH, Ong BK. Computer-based testing of the modified essay question: the Singapore experience. *Med Teach*. 2007 Nov; 29(9):e261-8.

6- Hammond EJ, McIndoe AK, Sansome AJ, Spargo PM. Multiple choice examination: Adapting an evidence –based approach to exam

[Downloaded from edujournal.zums.ac.ir on 2026-11-09]

[DOR: 20.1001.1.22519521.1396.10.27.6.2]

[DOI: 10.29252/edcj.10.27.19]

technique. *Anesthesia*. 1998; 53(11): 1105-1108.

7- Wadley M, Weaver SB, Curry C, Carthon C. Pharmacy students' perceptions of ExamSoft® as the primary assessment tool in an integrated therapeutics course. *Curr Pharm Teach Learn*. 2014; 6(6).

8- Morrow JA, Pulido T, Smith B, McDaniel TF, Willcox AB. Effective Use of e-Grading in the Dental Simulation Clinic. *J Dent Educ*. 2014; 78(6): 829-837.

9- Cantillon P, Irish B, Sales D. Using computers for assessment in medicine. *BMJ*. 2004; 329:606

10- Amin M M, Shayan S, Hashemi H, Poursafa P, Ebrahimi A. [Analysis of Multiple Choice Questions Based On Classical Test Theory]. *Iran J Med Educ*. 2011; 10 (5) :719-725. [Persian]

11- kazemi A, Ehsanpour S. [Item Analysis of Core Theoretical Courses Exams for Midwifery Students in Isfahan University of Medical Sciences]. *Iran J Med Educ*. 2011; 10 (5) :643-650. [Persian]

12- Shakurnia AH, Mozaffari AR, Khosravi Brougeni A. Survey on Structural of MCQs of residency exam in AJUMS. *Sci Med J.* 2010; 8(4):491- 502. [Persian]

13- Baig M, Ali SK, Ali S, Huda N. Evaluation of Multiple Choice and Short Essay Question items in Basic Medical Sciences. *Pak J Med Sci*. 2014; 30(1): 3–6.

14- Zolfaghari B, Adibi, Nasrin Farz, Ton Mozar, Karbasi, Strong P. Academic Achievement Tests. Isfahan: Neshat Newspaper, 2000: 59-107.

15- Zolfaghari B, Adibi, Nasrin Farz, Ton Mozar, Karbasi, Strong P. Academic Achievement Tests. Esfahan: Neshat Newspaper, 2000: 191-214.

16- Boevé AJ, Meijer RR, Albers CJ, Beetsma Y, Bosker RJ. Introducing Computer-Based Testing in High-Stakes Exams in Higher Education: Results of a Field Experiment. *PLoS One*. 2015 Dec 7; 10(12):e0143616.

17- Rodríguez-Díez MC, Alegre M, Díez N, Arbea L, Ferrer M. Technical flaws in multiple-choice questions in the access exam to medical specialties ("examen MIR") in Spain (2009-2013). *BMC Med Educ*. 2016 Feb 3; 16:47.

18- Tarrant M, Ware J. Impact of item-writing flaws in multiple-choice questions on student achievement in high stakes nursing assessment. *Med Educ*. 2008; 42(2):198-206.

19- Hochlehnert A, Brass K, Möltner A, Schultz JH, Norcini J, Tekian A, Jünger J. Good exams made easy: the item management system for multiple examination formats. *BMC Med Educ*. 2012 Aug 2; 12:63.

20- Shetty AS, Grajo JR, Decker S, Heitkamp DE, DeStigter KK, Mezwa DG, et. Al. ABR Core examination preparation: results of a survey of fourth-year radiology residents who took the 2013 examination. *Acad Radiol*. 2015 Jan; 22(1):121-9.

21- Evaniew N, Holt G, Kreuger S, Farrokhyar F, Petrisor B, Dore K, et.al. The orthopaedic in-training examination: perspectives of program directors and residents from the United States and Canada. *J Surg Educ*. 2013; 70(4):528-36.