



International Journal of Advanced Biotechnology and Research (IJBR)  
ISSN 0976-2612, Online ISSN 2278-599X,  
Vol-7, Special Issue-Number5, 2016, pp1674-1680  
<http://www.bipublication.com>

## Research Article

# A Comparison of Multiple-Choice and Essay Questions In The Evaluation of Dental Students

Shima Aalaei<sup>1</sup>, Mina Azizi Tas Ahmadi<sup>2</sup> and Azadeh Aalaei<sup>3\*</sup>

<sup>1</sup>Assistant Professor of Prosthodontics, Dental Caries Prevention Research Center,  
Qazvin University of Medical Sciences, Qazvin, Iran

<sup>2</sup>General dentist

<sup>3</sup>Master's Degree in Architecture, Islamic Azad University,  
Central Tehran Branch, Tehran, Iran

\*Corresponding author: Email: [azadehaalaei@gmail.com](mailto:azadehaalaei@gmail.com)  
Tel: +98-2833353064 ; Fax: +98-2833353066

## ABSTRACT

Evaluation of students is considered one of the most important aspects of teaching, and motivation for learning. This study compares the multiple-choice and essay questions in the evaluation of dental students. A total of 72 students participated in this study and were placed in two groups: Group A (N = 30), took an essay exam first and then took the same exam in multiple-choice format. Group B (N = 42), the sequence of exams was reversed. The students' grades, the difficulty and discrimination indices, and the students' responses to the two formats of the questions were analyzed using SPSS 21. *T*-test was used to compare the means. The students' grades in both groups were significantly higher on the multiple-choice exam than on the essay exam ( $13.5 \pm 2.3$  vs.  $12.7 \pm 2.8$  in Group A;  $12.7 \pm 3.3$  vs.  $12 \pm 2.5$  in Group B). The differences of discrimination index between two types of exams were significant (Group A:  $0.61 \pm 0.025$  vs.  $0.15 \pm 0.018$ ; Group B:  $-0.25 \pm 0.029$  vs.  $0.23 \pm 0.024$ ) unlike the difficulty index (Group A:  $0.74 \pm 0.012$  vs.  $0.745 \pm 0.012$ ; Group B:  $0.70 \pm 0.023$  vs.  $0.67 \pm 0.021$ ). Moreover, the students' responses to each question in both formats was the same in 67% of the cases in Group A and 66% of the cases in Group B. This study indicates that the multiple-choice questions were easier to answer than the essay questions and required less study effort.

**Keywords:** Multiple-choice, Essay questions, Difficulty index, Discrimination index, Exam

## [1] INTRODUCTION

Assessment or examination is defined as a systematic process of collecting, analyzing, and interpreting data in order to validate or judge students. Evaluation of students is considered one of the most important aspects of teaching, and its correct implementation induces motivation for learning and provides educational feedback for teachers [1]. Evaluation is carried out continuously during and at the end of education so that the weak and strong points of the teaching methods or educational programs can be

determined. Thus, accurate analysis of examinations is absolutely necessary for the promotion of the quality of educational systems [2,3].

Two types of questions that are commonly used for the evaluation of students are multiple-choice and essay questions. Multiple-choice questions are an example of objective questions, which determine the intended responses by the students. These exams take a short time to correct and make it possible to eliminate the personal role of the

teacher in the correction process. In large-scale settings where there are a large number of participants (e.g., university entrance examinations), multiple-choice exams are inevitable [4,5].

Essay exams, on the other hand, require a lot of time to correct because the examinees present their knowledge in an essay format. Therefore, such exams are most often used in small-scale settings such as educational courses where the number of examinees is limited [6].

It is still unclear which question type best shows the knowledge level of students in medical fields who directly deal with the health of the people. Hence, it is necessary to evaluate both assessment techniques in relation to the learning level of dental students as this can increase educational productivity. Evidence shows that the type and quality of exams significantly affect the educational quality and the teaching/learning process [7].

Questions are usually evaluated using the two criteria of difficulty index and discrimination index. The difficulty index expresses the percentage of individuals who have provided a correct answer to a question. This index is a number between 0 and 1, with the former indicating that no one has provided a correct answer and the latter showing that all the students have answered the question correctly. In other words, values near 0 show that the question was difficult and values near 1 indicate that the question was easy [8].

The discrimination index determines the ability of a question to distinguish strong students from weak ones. This index is a value ranging from -1 to 1, with values closer to 1 indicating a higher discrimination capacity. On the other hand, negative values indicate that the students who got low grades on the exam provided better responses than the student who earned good grades [8].

Walstad *et al.* evaluated differences between multiple-choice and essay questions in economics programs and reported higher grades for students on multiple-choice questions than on essay

questions. They argued that students can use their skills and the depth of their knowledge to answer essay questions [9].

Fathabadi and Seif showed that multiple-choice questions cause the students to study superficially while essay questions prompt students to study profoundly [10].

Mirfeiziet *al.* performed a study to evaluate third-year midwifery students' knowledge and clinical skills in pathophysiology, internal diseases, infectious diseases, and surgery using the OSCE (objective structured clinical examination) technique, conventional approaches, and the tests carried out in relation to practical subjects. They concluded that the OSCE is a good technique for the evaluation of the clinical skills of midwifery students [11].

Oppong evaluated the performance of male and female students of history using multiple-choice and essay questions and reported that while essay questions were more efficacious, multiple-choice questions were faster and easier [12].

The aim of the present study was to compare multiple-choice and essay questions (as two different testing techniques of evaluating dental students) drawing upon the difficulty and discrimination indices.

## **[II] MATERIALS AND METHODS**

This study has been reviewed by the medical research ethical committee of the Qazvin University of Medical Sciences and there is no conflict with ethical considerations. In the present descriptive/analytical crossover study, multiple-choice and essay tests (on the subject of theoretical fixed prosthodontics) were administered to two groups (A and B) of students. Due to the limited number of students, the tests for Group A (students entering the university in 2009; N = 30) were administered in February 2014, and the tests for Group B (students matriculating in 2010; N = 42) were administered in February 2015.

A total of 18 questions were administered to each group, making a total of 36 questions. In Group A,

the students first answered the essay questions and then answered the same questions in multiple-choice format. In Group B, the students first answered multiple-choice questions and then answered the same questions in essay format. All exams were administered under the supervision of the professors of the Department of Prosthodontics as semester-final exams.

This particular procedure was adopted in the knowledge that on the one hand familiarity with correct answers after doing the multiple-choice

questions can affect the answers to the essay questions, and that on the other hand concentration and thinking about the essay questions can affect the answers to the multiple-choice questions. Thus, the students' replies to both formats of the questions were compared.

Once the exam papers were corrected, the questions were analyzed in order to determine the difficulty and discrimination indices. The following formulas were used to calculate the two indices:

#### *Difficulty index*

$$= \frac{\text{correct response in the low - grade group} + \text{correct response in the high - grade group}}{\text{the number of students in the low - grade group} + \text{the number of students in the high - grade group}}$$

$$= \frac{\text{correct response in the high - grade group} - \text{correct response in the low - grade group}}{\text{the number of students in one group}}$$

*Discrimination index*

The examinees whose grade was 12 and above (out of 18) were considered the high-grade group, and those who scored below 12 were regarded as the low-grade group.

The values pertinent to the difficulty index were interpreted as follows: 0-0.2, extremely difficult; 0.2-0.4, very difficult; 0.4-0.6, difficult; 0.6-0.8, moderate; and 0.8-1, easy [8].

The percentages pertinent to the discrimination index were interpreted as follows: 0-20%, negative; 21-40%, poor; 41-60%, moderate; 61-80%, good; and 81-100%, very good [8].

To compare the students' responses to the two formats of the questions, the questions were divided up into four groups based on the responses: (1) correct response on both multiple-choice and essay exams; (2) correct response on the multiple-choice exam but incorrect response on the essay exam; (3) incorrect response on both exams; and (4) incorrect response on the multiple-choice exam but correct response on the essay exam. This classification gave us 2,592 responses: 72 students providing answers to 18 multiple-choice questions and also as many essay questions. (72×18×2=2592)

#### **2.1. Statistical analysis**

Mean, minimum, maximum, and standard deviation were used as descriptive parameters.

The paired-sample *t*-test was used to compare the multiple-choice and essay exams in terms of mean grades, difficulty index, and discrimination index. Since there were different numbers of students in the two groups, the percentages were used to compare the performance of the two groups. Statistical significance was set at  $\alpha=0.05$ . Data management and analysis was performed using SPSS 21 (IBM Corporation, USA, 2012).

#### **2.2. Ethical considerations**

Only the students who were willing to participate in the study did the exams. Moreover, the grades of the students were kept confidential, and the authorities of the Faculty were only given the results of the study.

### **[III] RESULTS**

The students in Group A scored significantly differently on the multiple-choice (Mean: 13.5) and essay (Mean: 12.7) exams ( $P=0.011$ ). A similar difference was observed for the students in Group B: a mean of 12.7 on the multiple-choice exam and 12 on the essay exam ( $P=0.017$ ).

Table 1 and Table 2 present the frequency distributions of the difficulty and discrimination indices obtained for the multiple-choice and essay questions answered by the students in both study groups.

In both groups (Table 3), no significant difference was found between multiple-choice and essay questions in terms of the means of difficulty

indices (Group A:  $0.74 \pm 0.012$  vs.  $0.745 \pm 0.012$ ; Group B:  $0.69 \pm 0.023$  vs.  $0.68 \pm 0.021$ ) ( $P > 0.05$ ).

**Table 1.** The frequency distribution of the difficulty indices obtained for the multiple-choice and essay questions

Classification	Group A Multiple-choice	Group B Multiple-choice	Group A Essay	Group B Essay
	Frequency (%)	Frequency (%)	Frequency (%)	Frequency (%)
Extremely difficult	0 (0)	1 (5.55)	0 (0)	0 (0)
Very difficult	0 (0)	2 (11.11)	0 (0)	5 (27.77)
Difficult	6 (33.33)	5 (27.77)	6 (33.33)	3 (16.66)
Moderate	8 (44.44)	6 (33.33)	8 (44.44)	6 (33.33)
Easy	4 (22.22)	4 (22.22)	4 (22.22)	4 (22.22)
Total	18 (100)	18 (100)	18 (100)	18 (100)

**Table 2.** The frequency distribution of the discrimination indices obtained for the multiple-choice and essay questions

Classification	Group A Multiple-choice	Group B Multiple-choice	Group A Essay	Group B Essay
	Frequency (%)	Frequency (%)	Frequency (%)	Frequency (%)
Negative	2 (11.11)	2 (11.11)	1 (5.55)	17 (94.44)
Poor	8 (44.44)	5 (27.77)	0 (0)	1 (5.55)
Moderate	4 (22.22)	4 (22.22)	1 (5.55)	0 (0)
Good	4 (22.22)	7 (38.88)	16 (88.88)	0 (0)
Total	18 (100)	18 (100)	18 (100)	18 (100)

**Table 3.** A comparison of the means of difficulty indices obtained for the multiple-choice and essay exam in Group A and B

Study Group	Exam type	Mean	Mean difference	T	P-value
A	Multiple-choice	0.743	0.002	0.133	0.895
	Essay	0.745			
B	Multiple-choice	0.696	0.028	0.022	0.983
	Essay	0.668			

In contrast, in both groups of the study (Table 4), a significant difference was observed between multiple-choice and essay questions in terms of the means of discrimination indices.

As can be seen, in Group A, the discrimination index of the essay questions was more than that of the multiple-choice questions ( $0.61 \pm 0.025$  vs.  $0.15 \pm 0.018$ ) ( $P < 0.05$ ), but in Group B, the discrimination index of the essay questions was negative and lower than the multiple-choice questions ( $-0.25 \pm 0.029$  vs.  $0.23 \pm 0.024$ ) ( $P < 0.05$ ).

**Table 4.** A comparison of the means of discrimination indices obtained for the multiple-choice and essay exam in Group A and B

Study Group	Exam type	Mean	Mean difference	T	P-value
A	Multiple-choice	0.147	0.46	6.36	0.001
	Essay	0.607			
B	Multiple-choice	0.227	0.48	6.42	0.002
	Essay	-0.253			

An issue that is worthy of mention here is that the multiple-choice and essay tests administered to Group B proved insignificantly more difficult than their counterparts given to the students in Group A ( $P=24$  and  $21$  for the multiple-choice and essay tests, respectively).

As for the discrimination index, the values for the multiple-choice tests were almost the same in the two groups ( $P = 0.26$ ). However, the discrimination index of the essay test given to Group B was significantly lower than that of the essay test administered to Group A ( $P = 0.00$ ).

Finally, in order to more accurately assess multiple-choice and essay and questions, the response each student provided to each multiple-choice question was compared with the response that student gave to the same question in essay format (Table 5). As each question had been presented in both formats, it was expected that the responses would be the same or minimally different. Such an expectation, however, was not fulfilled, perhaps due to differences in the nature of multiple-choice and essay questions. More specifically, in Group A, the answers to 33% of the questions were not the same; this figure was 34% in Group B. In other words, the answer to a question in multiple-choice format was correct, but the answer to the same question in essay format was incorrect, and vice versa.

**Table 5.** A comparison of the responses provided by the students to the multiple-choice questions and their essay counterparts

	Group A		Group B	
	Frequency	Percentage	Frequency	Percentage
Correct multiple-choice / correct essay	309	57.22	337	44.57
Correct multiple-choice / incorrect essay	94	30	172	22.75
Incorrect multiple-choice / incorrect essay	90	16.66	157	20.76
Incorrect multiple-choice / correct essay	47	8.70	90	11.90

Table 6 presents a comparison of the questions which were answered differently in the multiple-choice and essay formats. Clearly, in both groups, the “correct multiple-choice / incorrect essay” case exhibited a higher percentage than the “incorrect multiple-choice / correct essay” case.

**Table 6.** A comparison of the questions which were answered differently in the multiple-choice and essay formats

	Incorrect multiple-choice / correct essay		Correct multiple-choice / incorrect essay		Total number of questions answered differently	
	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Group A	47	33%	94	67%	141	100
Group B	90	34%	172	66%	262	100

#### [IV]DISCUSSION

The main aim of the present study was to compare two different examination types (i.e., multiple-choice and essay questions) in the evaluation of dental students in terms of grades, the similarity or difference between the responses to the two formats of the same questions, and the difficulty and discrimination indices. The results showed higher grades on the multiple-choice questions compared to the essay questions, which is consistent with almost all previous studies. For instance, Oyebola *et al.* showed that the students' grades on multiple-choice questions were higher than on essay questions.

They also showed that the students' grades are correlated with the type of the exam used [13]. The results of the present study are also consistent with those reported in a study by Wilkinson *et al.* on different types of multiple-choice and essay questions for medical students [14].

In addition, Pepple *et al.* evaluated the performance of students in answering multiple-choice and essay questions and reported that the mean correct responses were 64% and 47%, respectively, indicating a significant difference between the two exam types [15].

Moreover, Delaram and Sharifi reported a higher mean grade on multiple-choice questions than on essay questions and concluded that the use of multiple-choice questions for the evaluation of students might be particularly helpful to the weaker students [16].

Sharma *et al.* performed a study to compare the medical students' performance on multiple-choice and essay questions. The results showed mean grades of 55.24% and 47.63%, respectively, indicating significantly higher grades on multiple-choice questions [17].

Another finding of the present study was the lack of a significant difference between the multiple-choice and essay questions in terms of the difficulty index although the exams given to Group B proved more difficult than those given to Group A.

Moreover, in Group A, the mean discrimination index of essay questions was significantly higher than that of multiple-choice questions; however, in Group B, the mean discrimination index of the essay questions was very low and negative.

This is attributable to the fact that knowledge of correct answers to the multiple-choice questions which were tried first served as a confounding factor, but the thinking and concentration applied to the essay questions attempted first did not have a significant effect on the multiple-choice questions which were answered later. Therefore, the discrimination indices of the multiple-choice questions in both groups were almost the same, but this was not the case in the discrimination indices of the essay questions in the two groups.

Lastly, a comparison of the responses provided by the students to the multiple-choice questions and their essay counterparts showed that in both groups of the study, the case where the multiple-choice questions were answered correctly but the essay questions received incorrect responses showed a higher percentage than the case where the essay questions were answered correctly but the multiple-choice questions were given incorrect answers.

A conclusion which can be drawn from this observation is that the students participating in the study found the multiple-choice questions easier to do than the essay questions, and this seems rather logical considering the fact that the multiple-choice questions generally require less study effort and can be answered correctly by chance.

#### [V] CONCLUSION

Regardless of the data obtained from Group B, which were inconclusive mainly in consequence of the effect of the moderating variable, the results of the present study showed that the students got significantly higher grades on the multiple-choice questions than on the essay questions.

Since the discrimination index of the multiple-choice questions was significantly lower than that of the essay questions in Group A, it appears that

essay questions provide more accurate evaluations compared to multiple-choice questions, and it is suggested that students be evaluated using essay questions more frequently.

## FINANCIAL DISCLOSURE

Nothing to declare

## REFERENCES

1. Haigh C (2003), An evaluation of a judgmental model of assessment for assessing clinical skills in MSc students. *Nurse Educ Pract.* 3:43-48.
2. Pepple DJ, Young LE, Carroll RG (2010), A comparison of student performance in multiple-choice and long essay questions in the MBBS stage I physiology examination at the University of the West Indies (Mona Campus). *Adv Physiol Educ.* 34:86-89.
3. Van Bruggen L, Manrique-van Woudenberg M, Spierenburg E, Vos J (2012), Preferred question types for computer-based assessment of clinical reasoning: a literature study. *Perspect Med Educ.* 1:162-171.
4. Karl M, Graef F, Eitner S, Beck N, Wichmann M, Holst S (2007), Comparison between computer-aided testing and traditional multiple choice: and equivalence stud. *Eur J Dent Educ.* 11:38-41.
5. Bauer D, Kopp V, Fischer MR (2007), Answer changing in multiple choice assessment change that answer when in doubt--and spread the word! *BMC Med Educ.* 7:28.
6. Nnodim JO (1992), Multiple-choice testing in anatomy. *Med Educ.* 26:301-309.
7. Tahernejhad K, Javidan F (2008), Advanced assessment of medical students' clinical performance: challenges, methods and approaches. *Strides in Development of Medical Education.* 5:58-70.
8. Sim DS, Rasiah RJ (2006), Relationship between item difficulty and discrimination indices in true/false-type multiple choice questions of a para-clinical multidisciplinary paper. *Ann Acad Med Singapore.* 35:67-71.
9. Walstad WB, Becker WE (1994), Achievement Differences on Multiple-Choice and Essay Tests in Economics. *American Economic Review.* 84:193-196.
10. Epstein RM, Hundert EM (2002), Defining and assessing professional competence. *JAMA.* 287:226-235.
11. Mirfeizi M, Mehdizadeh-Tourzani Z, Bahrami-Babaheydari T, Mirfeizi SZ, Asghari Jafarabadi M, Mirheidar M, et al. (2012), A Comparative Study on Evaluating of Knowledge and clinical Practices of Midwifery Students in the Courses of Physiopathology, Infectious and Internal Diseases and Surgery based on Objective Structured Clinical Examination and Traditional Assessment Methods. *Future of Medical Education Journal.* 2:3-8.
12. Oppong CA (2013), Gender Differences In Students' Performance In History At Senior High Schools In Cape Coast. *Journal of Arts and Humanities.* 2:34-39.
13. Oyebola DD, Adewoye OE, Iyaniwura JO, Alada AR, Fasanmade AA, Raji Y (2000), A comparative study of students' performance in preclinical physiology assessed by multiple choice and short essay questions. *Afr J Med Med Sci.* 29:201-205.
14. Wilkinson T, Hannah S (2015), Are Spot Test Multiple Choice Questions Easier to Answer than Short Answer Questions? *The FASEB Journal.* 29:344.1.
15. Pepple DJ, Young LE, Carroll RG (2010), A comparison of student performance in multiple-choice and long essay questions in the MBBS stage I physiology examination at the University of the West Indies (Mona Campus). *Adv Physiol Educ.* 34:86-89.
16. Delaram M, Sharifi A (2014), Comparison of Nursing students' grades on "Mothers' and Infants' Hygiene" exam in the essay and multiple-choice formats. *The Journal of Education Development in Medical Sciences.* 7:31-37.
17. Sharma HS, Mutalik MM (2014), Comparison between student scores in multiple choice questions and structured essay questions at iii mbbs examination in otorhinolaryngology in a medical college in Mauritius. *Int J Pharm Bio Sci.* 5:(B) 1195-202.