

Analysis Of Testing With Multiple Choice versus Fill In The Blank Questions: Outcome-Based Observations In Dental Subjects

Research Article

Sheeja S Varghese^{1*}, Asha Ramesh²

¹ Professor and Dean Department of Periodontics Saveetha Dental College, Saveetha Institute of Medical and Technical sciences Saveetha University, Chennai, India.

² Senior Lecturer Saveetha Dental College, Saveetha Institute of medical and technical sciences Saveetha University, Chennai, India.

Abstract

Background: The testing methods have evolved over time, but in dental education, the multiple-choice questions (MCQ) remain preferred among educators. There is limited research on Fill-in-the-blanks (FIB) that tap the student's 'deeper' learning potential. This retrospective study aims to assess the difference in learning outcomes of dental undergraduate students with MCQ and FIB-based assessments.

Methods: The final year students of different academic years were divided into two groups based on the assessment method in 'end-of-year' written examination (summative assessment) conducted on eight dental subjects: Group I students (n=75) were given MCQs, short, and long essays where as, for Group II students (n=75), FIB replaced the MCQ. The summative marks were compared between the two groups. Also, Group II students responded to an online questionnaire to assess the learning outcomes in their pre-final (MCQ-based) and final year (FIB-based) examinations.

Results: Independent t-test compared the summative assessment scores between the two groups and observed that there was significant difference favoring Group I in five subjects. Two subjects showed significant difference in scores favoring Group II and there was no difference in one subject. Also, majority of the students gave feedback that they had better learning outcomes with FIB format.

Discussion: The study concluded that there was a significant difference in the 'end-of-year-written examination scores between the two groups and the academic performance was better with MCQ format. But, the questionnaire feedback from the students observed that FIB format facilitated in-depth learning and comprehension of the subject.

Keywords: Assessment Methods; Learning Outcome; Multiple Choice Question; Fill In The Blank; Dental Education.

Background

In dental education, the art and science behind learning has undergone a paradigm shift with remarkable progress. Competency-based education is the need-of-the-hour as it orients the subject to their outcome abilities and organizes around their competencies [1]. The effective delivery of healthcare not only banks on theoretical knowledge and technical skills, but also on interpersonal and analytical skills, interdisciplinary care, and following an evidence-based approach [2]. The valid assessment of clinical

competence is the actual performance of the doctor in a clinical setting [3]. This asserts that our assessment systems should be more sound, and robust to comprehensively evaluate the required attributes along with the essential knowledge and skill-set.

Tests are exclusively used assessment tools in medical and dental education. They are used in assigning grades and to ratify professional competence. There are two classes of memory tests and they include: i) recognition tests that employ the selection of appropriate response from a list of alternatives (Multiple choice

*Corresponding Author:

Sheeja S Varghese,

Professor and Dean Department of Periodontics Saveetha Dental College, Saveetha Institute of Medical and Technical sciences Saveetha University, Chennai, India.

Tel: 9884042252

Email Id: sheejavarghese@saveetha.com

Received: March 05, 2021

Accepted: March 12, 2021

Published: March 17, 2021

Citation: Sheeja S Varghese, Asha Ramesh. Analysis Of Testing With Multiple Choice versus Fill In The Blank Questions: Outcome-Based Observations In Dental Subjects. *Int J Dentistry Oral Sci.* 2021;08(03):2020-2024. doi: <http://dx.doi.org/10.19070/2377-8075-21000397>

Copyright: Sheeja S Varghese ©2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

questions (MCQs), true/false) and, ii) production tests that mandates the student to generate the best response to the question (fill-in-the-blanks (FIB, essay) [4]. It has been shown extensively in research that production tests involve more effortful retrieval of information to generate a response and that results in better retention when compared to recognition tests [5]. The educational context of assessment methods has been an understudied subject of research and the importance of this topic can be condensed in the aphorism that assessment drives learning [4].

Globally, selected response questions or MCQs have been a popular method for summative assessment of medical knowledge [6]. The students are required to select the single best response among three or more options. They are specifically favored for its many advantages like high reliability and validity, high objectivity, easy evaluation of the answer scripts, and for its utility as a tool for self-assessment [7]. In the current literature, various shortcomings of the MCQs have been highlighted and they include: i) Can only assess the lower-order thinking such as memorizing facts and the ability to recall factual information, ii) Guessing the answer that can reduce the reliability of the assessment results, iii) Cannot be used to assess the practical aptitude such as patient communication skills [3, 7, 8].

'FIB' or 'short answer questions' are open-ended and the students are required to frame answers that are no more than one or two words. They are similar to multiple-choice questions in its construction [3]. In contrast to the MCQs, the need to select an answer from a fixed number of options and also guessing the answer can be completely eliminated. From a list of four options, the student has 20% chance of selecting the correct answer based on guess work in MCQs, but there is no such practice in FIB because the answer is a single best response to the framed question [9]. In medical education research literature, a fund of clinical knowledge that is segregated into usable networks is essential for clinical practice and problem solving, and studies are required in this area to evaluate the effectiveness of these assessment methods [4].

Feedback is an integral element of assessment and the methods of assessment should inform the learners about their progress towards becoming experts [10]. Van der Vleuten et al stated that the best assessment practice must provide an opportunity for formative feedback that aids in improved performance [11, 12]. Constructive student feedback is essential for competency-based education and it guides the student as well as the educator towards measuring progress in acquiring core knowledge and competencies.

Assessment can drive learning and learning behaviors, determine learning objectives, and can modify the curriculum itself. A strong assessment system can motivate learning, instill values, and strengthen competence [13]. The employment of a valid assessment tool is crucial in shaping the professional development of the students as they require an amalgamation of theoretical knowledge and clinical acumen to be incorporated in their practice. Even though there are many studies and reviews elaborating on the utility and reliability of these assessment methods in medical education, there is a dearth in studies pertaining to the field of dentistry and the shaping of the future of dental education is reliant on the credibility of these assessment methods. Thus, the aim of this retrospective study was to compare the influence

of different assessment methods on the learning outcome of students in terms of academic performance spanning over two academic years.

Methods

This retrospective study was performed in a dental school in India and it included 150 dental undergraduate students who were pursuing the final year of their studies. This study was approved by the Institutional Review Board (SRB/FACULTY/05/18/PERIO/08). Group I students (n=75) belonged to the academic year of 2016-17, where as, Group II students (n=75) belonged to 2017-18 period. For Group I students, the end-of-course written examination (summative assessment) comprised of MCQs, short, and long essays whereas, for Group II students, FIB, short, and long essays were utilized for the assessments. For Group II students, The University 'Academic Council' introduced the FIB format as part of the examination reforms and it was intimated to the students prior to the course commencement. Both the groups appeared for a three-hour written examination in eight subjects covering the different specialties in dentistry. There were no syllabus changes, no differences in teaching methodologies, and no change in the number of contact hours for the two groups of students, there by, ascertaining the standardization of the teaching-protocol. The short and long essay questions were chosen from a standardized question bank whose contents were categorized according to a difficulty level and the University administered 'end-of-course' examinations were similar for both the groups.

Following the final year written examination, the Group II students were given a feedback questionnaire containing five items. The Group II students were chosen for this feedback because their pre-final year exams followed the MCQ pattern, where as the final year exams contained FIB. The questionnaire items aimed to assess the differences observed by the students leading to the preparation of MCQs and FIB for the respective exams and the questions were pertaining to the anxiety level during exam preparation, thorough subject knowledge obtained after the exams, confidence levels, and future utility of their preparation for postgraduate qualifying examinations.

The obtained results were subjected to the Kolmogorov-Smirnov and Shapiro-Wilks tests of normality. The resultant data showed that they followed a parametric distribution. Independent t-test was performed to compare the average scores obtained in the summative assessment between the two groups. All analyses were conducted using statistical software (SPSS software, version 17). $p < 0.05$ was considered to be statistically significant. A descriptive analysis was performed for the questionnaire items.

Results

A retrospective study included 150 dental undergraduate students (Group I=75; Group II=75) and compared the learning outcomes of students based on summative assessment marks in the end-of-course final year examination covering eight subjects in the different fields of dentistry. The students were divided into two groups based on the assessment methods and different comparison was based on assessment tools that were employed in the examination pattern like MCQ and FIB. Independent t-test determined that there was a statistically significant difference among

the two groups in seven subjects. Summative marks were significantly higher for Group I students in subjects like Periodontics (p value: 0.000), Oral and Maxillofacial Surgery (p value: 0.002), Oral Medicine (p value: 0.000), Prosthodontics (p value: 0.000), and Pediatric Dentistry (p value: 0.025). There was no statistically significant difference in summative marks in the Conservative Dentistry and Endodontics subject (p value: 0.686) where as, it was significantly higher favoring Group II students in the subjects of Public Health Dentistry (p value: 0.000) and Orthodontics (p value: 0.000) (Table 1).

An online feedback questionnaire was given to Group II students to evaluate the differences in their perspective, learning outcome, and study preparation to different assessment methods like MCQ

and FIB. There were 71 respondents to the questionnaire, of which 53.5% of the students observed that they were apprehensive about FIB during the exam preparation and after the exams (Figure 1), 69% of the students felt that they had read the subjects thoroughly while preparing for FIB questions (Figure 2). To further assess the utility of these assessment tools for future competitive exam preparation, 63.4% of students observed that FIB preparation would help them for future examinations (Figure 3). The confidence level of the students on the theoretical knowledge after their pre-final (MCQ pattern) and final year (FIB pattern) exams were ranked between 1-10 and the MCQ's observed a highest score of 6 (28.2%), where as, FIB derived a highest score of 8 (32.4%) (Figure 4, 5).

Table 1. Independent t-test for comparing summative assessment marks scored in eight dental specialty subjects between the two groups.

Variable	Subjects	Groups	n	Mean±SD	t-value	p-value
Summative Marks	Oral Surgery	Group I	75	70.01±3.57	3.235	0.002
		Group II	75	66.72±8.03		
	Periodontics	Group I	75	73.03±2.87	6.485	0.000
		Group II	75	68.00±6.07		
	Conservative Dentistry	Group I	75	68.68±3.20	0.405	0.686
		Group II	75	68.39±5.38		
	Orthodontics	Group I	75	64.36±5.17	-6.856	0.000
		Group II	75	71.65±7.62		
	Prosthodontics	Group I	75	75.17±5.46	5.698	0.000
		Group II	75	70.28±5.04		
	Pediatric Dentistry	Group I	75	73.55±5.11	4.472	1.000
		Group II	75	69.88±4.94		
	Oral Medicine & Radiology	Group I	75	65.72±5.04	6.601	0.000
		Group II	75	67.76±5.93		
	Public Health Dentistry	Group I	75	71.64±4.99	-2.261	0.025
		Group II	75	64.90±7.29		

p-value<0.05 is considered to be statistically significant.

Figure 1. Showing the fill in the blank questions made more number of students apprehensive during exam preparation.

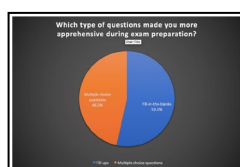


Figure 2. Shows that majority of students felt that fill in the blank questions made them read the subjects thoroughly.

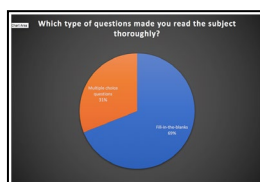


Figure 3. Shows that preparation for fill in the blank questions will be more helpful for future competitive examinations.

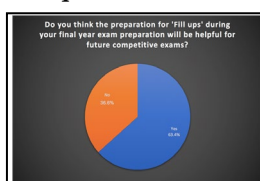


Figure 4. Shows confidence level of group I students where the majority of the students had the scores from 5-7.

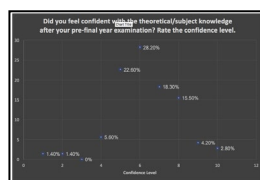
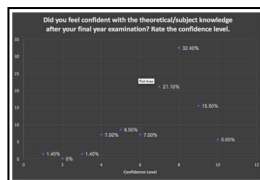


Figure 5. Showing the confidence level of group II students and the majority had the scores from 7-9.



Discussion

The results of this study are consistent with the notion that FIB or short answer questions are open-ended and require deeper learning strategies from the dental undergraduate students. It is conceived that student's preparation of course content can get altered based on the type of test they anticipate and this can influence the nature and quality of student learning [14]. Retrieval of information using short answer questions bring about the best learning due to the fact that there is increased retrieval effort or difficulty [15]. It can be observed that the 'end-of-year' summative marks were lower for Group II students who attempted FIB in five out of eight dental subjects. This negative outcome could be attributed to the fact that the students were exposed to this format of assessment only in their final year and it also corroborated with the result from the questionnaire that they were more apprehensive about FIB than MCQ during the exam preparation. Also, similar results were observed in a study conducted by MelovitzVasan et al., on first year medical students in a preparatory anatomy course. The modules were assessed using MCQ and open-ended question (OEQ) and the initial thorax module showed that OEQ marks were lower than that of MCQ marks. The study stated that it could also be possible that the students were learning through memorization or shallow processing and not getting actively engaged with the course material during the initial module [16].

This study observed that there was statistically significant difference favoring Group II students in two subjects (Orthodontics and Public Health Dentistry). Also, there was no significant difference in the scores between the two groups in the subject of Conservative Dentistry and Endodontics. The variability in part could be attributed to the course content of the different subjects and the students' inclination towards a specific subject. This can explain the better performance seen among students who appeared for the FIB format of examination. There has been a considerable difference in the results with students scoring higher with FIB in some subjects, better scores with MCQ in the remaining subjects, and similar scores in both the formats in one subject. It is reasonable to suppose that customizing assessment methods to individual subjects could be a potential thrust area of future research.

MCQ assessment does not tap the student's higher order of thinking and it can result in higher chances of guessing the correct answer and that in turn can reduce the reliability of this method for

students with lower ability [17]. It can also mislead the faculty and educators about the student's subject knowledge or proficiency in the course [18]. There are also numerous flaws that can be encountered in MCQ format where there are irrelevant difficulties in negative stem questions, ineffective distractors, and low discrimination index [19, 20]. MCQ format can aid in 'superficial learning' or 'surface learners' and this can be correlated with the results from the feedback where majority of the students had a lower confidence level in terms of subject knowledge after their pre-final year examination employing MCQ's. Also, in a recent study conducted by Sam et al, it was noted that students who were strategic and practiced large number of past questions could get adept at choosing the correct answer without having an in-depth comprehension of the subject [21].

The prominent results of the study were from the questionnaire feedback conducted on Group II students and it observed that there was an increased subject knowledge, higher confidence level, and definitive future utility of FIB preparation for competitive exams. The competencies required for a junior dentist is the ability to recall the correct diagnosis and formulate a treatment plan for the different case scenarios [22]. The attribute of FIB includes its testing of reasoning abilities like evaluation, analysis, and synthesis of knowledge in an abstract way without the 'cueing effect'[23]. Answering a short answer question necessitates active generation of an answer through recollection of memory or source text, where as, success in answering a MCQ can be partly based on familiarity [24]. It can be derived that preparation for FIB has a positive influence on the reader's active engagement with the course material, comprehension of the subject, and overall deeper learning of the content.

Although there are many advantages with FIB, the evaluation poses a difficulty wherein the faculty members should invest a significant amount of their time to read and evaluate the answers. Also, in our study it was observed that the evaluation of FIB questions was more time-consuming compared to MCQ format as reported by the examiners. The study was carried out real-time and not in a simulated environment but a small sample size could be a potential limitation in this study. Although there is a deficit of marks with the use of FIB when compared to MCQ, the onus on student learning and understanding of the subject was better focused with FIB format as observed from the positive feedback from the students in the questionnaire. Curriculum has become technology-driven and the students are better poised for challenges, so qualitative changes in the assessment methods can

improve student behavior and academic performances for achieving a competency-based education. A single mode of assessment cannot be generalized for all the subjects as the course content can have a difference in the theoretical and evidence-based clinical content, a customized assessment method for individual subjects can be the target of future dental-based educational research.

The study concluded that there was a significant difference in the end-of-year examination marks of dental undergraduate students in favor of MCQ when compared to FIB format. But, the results from the questionnaire feedback from the students suggested that FIB was the most preferred format for assessment. Although there are different assessment methods employed in dental education, a single ideal mode of assessment remains to be elusive. FIB or short answer questions inculcates active learning from the students and that leads to better comprehension of subject knowledge. Future examinations can incorporate FIB and validate its results with other tools of assessment, there by, adding new paradigm to assessments in dental education.

References

- [1]. Frank JR, Mungroo R, Ahmad Y, Wang M, De Rossi S, Horsley T. Toward a definition of competency-based education in medicine: a systematic review of published definitions. *Med Teach.* 2010;32(8):631–7. PubmedPMID: 20662573.
- [2]. Tabish SA. Assessment methods in medical education. *Int J Health Sci.* 2008 Jul;2(2):3–7. Pubmed PMID: 21475483.
- [3]. Al-Wardy NM. Assessment methods in undergraduate medical education. *Sultan Qaboos Univ Med J.* 2010 Aug;10(2):203–9. PubmedPMID: 21509230.
- [4]. Larsen DP, Butler AC, Roediger III HL. Test-enhanced learning in medical education. *Med Educ.* 2008 Oct;42(10):959–66. PubmedPMID: 18823514.
- [5]. McDaniel MA, Roediger HL, McDermott KB. Generalizing test-enhanced learning from the laboratory to the classroom. *Psychon Bull Rev.* 2007 Apr;14(2):200–6. Pubmed PMID: 17694901.
- [6]. Schuwirth LWT, van der Vleuten CPM. ABC of learning and teaching in medicine: Written assessment. *BMJ.* 2003 Mar 22;326(7390):643–5. Pubmed PMID: 12649242.
- [7]. Nnodim JO. Multiple-choice testing in anatomy. *Med Educ.* 1992 Jul;26(4):301–9. PubmedPMID: 1630332.
- [8]. Kołodziejczak B, Roszak M, Ren-Kurc A, Kowalewski W, Bręborowicz A. E-assessment in medical education. 2016.
- [9]. Chandratilake M, Davis M, Ponnampereuma G. Assessment of medical knowledge: the pros and cons of using true/false multiple choice questions. *Natl Med J India.* 2011 Aug;24(4):225–8. PubmedPMID: 22208143.
- [10]. Humphrey-Murto S, Wood TJ, Ross S, Tavares W, Kvern B, Sidhu R, et al. Assessment Pearls for Competency-Based Medical Education. *J Grad Med Educ.* 2017 Dec;9(6):688–91. Pubmed PMID: 29270255.
- [11]. Van Der Vleuten CPM, Schuwirth LWT, Driessen EW, Govaerts MJB, Heeneman S. Twelve Tips for programmatic assessment. *Med Teach.* 2015 Jul;37(7):641–6. Pubmed PMID: 25410481.
- [12]. van der Vleuten CPM, Schuwirth LWT, Driessen EW, Dijkstra J, Tigelaar D, Baartman LKJ, et al. A model for programmatic assessment fit for purpose. *Med Teach.* 2012;34(3):205–14. PubmedPMID: 22364452.
- [13]. Case S, Swanson D. Constructing written test questions for the basic and clinical sciences. 3rd Ed. Philadelphia: National Board of Medical Examiners; 2002. 180 p.
- [14]. Skimin M, Kuechler W. Multiple-choice tests and student understanding: What is the connection? *Journal of Innovative Education.* 2005 Jan;3(1):73–98. Pubmed PMID: 33091384.
- [15]. Smith MA, Karpicke JD. Retrieval practice with short-answer, multiple-choice, and hybrid tests. *Mem Hove Engl.* 2014;22(7):784–802. PubmedPMID: 24059563.
- [16]. Melovitz Vasan CA, DeFouw DO, Holland BK, Vasan NS. Analysis of testing with multiple choice versus open-ended questions: Outcome-based observations in an anatomy course. *Anat Sci Educ.* 2018 May 6;11(3):254–61. PubmedPMID:28941215.
- [17]. Cronbach L. Five perspectives on validity argument. In: Wainer H, Braun HI (Editors). *Test validity*, 1st Ed. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.; 1988. 3-17 p.
- [18]. Chan N, Kennedy P. Are multiple-choice exams easier for economics students. *Southern Economic Journal.* 2002;68(4).
- [19]. Rodríguez-Díez MC, Alegre M, Díez N, Arba 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. PubmedPMID: 26846526.
- [20]. Kaur M, Singla S, Mahajan R. Item analysis of in use multiple choice questions in pharmacology. *Int J Appl Basic Med Res.* 2016 Sep;6(3):170–3. PubmedPMID:27563581.
- [21]. Sam AH, Hameed S, Harris J, Meeran K. Validity of very short answer versus single best answer questions for undergraduate assessment. *BMC Med Educ.* 2016 Oct 13;16(1):266. PubmedPMID:27737661.
- [22]. Geng Y, Zhao L, Wang Y, Jiang Y, Meng K, Zheng D. Competency model for dentists in China: Results of a Delphi study. *PloS One.* 2018;13(3):e0194411. PubmedPMID:29566048.
- [23]. Stiggins R. *Student-Involved Assessment for Learning*. 4th Ed. Upper Saddle River, NJ: Pearson Education; 2005. 400 p.
- [24]. Ozuru Y, Briner S, Kurby CA, McNamara DS. Comparing comprehension measured by multiple-choice and open-ended questions. *Can J Exp Psychol Rev Can Psychol Exp.* 2013 Sep;67(3):215–27. PubmedPMID:24041303.