ORIGINAL ARTICLE

Quantification of the Content Validity of Written Tests at Mashhad Dental School in 2011-2012

Javad Sarabadani[1]; Majid Sanatkhani[1]; Maryam Amirchaghmaghi [1,*]; Majid Fehresti Sani[2]

[1]Oral and Maxillofacial Diseases Research Center, Department of Oral Medicine, School of Dentistry, Mashhad University of Medical Sciences, Mashhad, IRAN
[2]General Dentist

* Vakilabad Blvd, Mashhad Dental School
Mashhad, IRAN

Tel: 09155164476
Fax: 009851 38829500
Email:
amirchakhmaghim@mums.ac.ir

Background: Academic tests are valuable tools for evaluating students' knowledge and capabilities. This study aimed to assess the content validity of final examinations at Mashhad Dental School. The practical purpose of this study was to enhance the quality of written examinations in this school.

Methods: The present descriptive, cross-sectional study aimed to determine the content validity of final examination of Mashhad Dental School, located in Northeast of Iran during 2011-2012. In this study, in collaboration with the Education Development Office (EDO) of Dental School, we designed this study to assess the content validity of written tests in the school. The percentage of content coverage (CC) and content relevance (CR) of each item were determined, and the percentage of content validity (CV) was calculated using Messick method. We used frequency distribution tables and diagrams in order to assess the content validity of final examinations.

Results: In total, 68 faculty members who were designers of exam questions participated in the current study. The mean percentage of CR and CC were 95.2% and 76.87 in the final exams of all departments, respectively. In addition, the mean percentage of content validity was 86.04 in the exams of different groups.

Conclusions: This study concluded that the mean percentage of CR and CC of final exams was good (95.2% and 76.87%) in all educational but in many groups, it was not desirable. Educational seminars for awareness of teachers from the importance of content validity and quantification method are quite effective in designing efficient exams with greater sensitivity.

Keywords: Written Test; Content Validity; Evaluation; Dentistry

تعيين مستوى الدقة في الاختبارات الكتبية التي تم اجراؤها في كلية طب الاسنان في مدينة مشهد خلال العام الدراسي ٩٠-٩١ هـ.ش .

المقدمة: إن مستوى الدقة هو اصطلاح يشير الى الهدف الذي برمو اليه صنع ذلك الاختبار. إن الاختبارات الجامعية تعتبر ادوات مهمة في تقييم التعلم و القدرة عند الطلاب. لذا من الضروري متابعة مستوى الدقة والكفاء في هذه الاختبارات. إن الهدف من هذه الدراسه هو تعيين مستوى الدقة لمحتوى الاختبارات الكتبية التي اجريت في كلية طب الاسنان في جامعه مشهد في الفصل الاول من العام الدراسي ٩٠-٩١ هـ.ش .

الأساليب: في هذه الدراسة التوصيفية و المقطعية التي تم اجراؤها بمساعده مكتب التوسعه التعليمية (EDO)في كلية طب الاسنان في جامعه مشهد تم استخدام أسلوب (مسيك) لتعيين مستوى الدقة في محتوى الاختبارات. لذلك تم اختيار اختبار نهائي من كل مجموعه تعليميه في الفصل الاول من العام الدراسي ٩٠-٩١ بشكل عشوائي. و بعد تم الحصول على جدول(Blue Print) لذلك الاختبارمن خلال استاذ في المجموعة التعليمية و في النهايه تم تعيين مستويات التناسب المحتوائي (CR) و النقطة المحتوائي (CC)و الدقة المحتوائيه (CV) من خلال الاساتذه الذي وضعوا هذه الاختبارات.

النتائج: اشترك في هذه الدراسه ١١ مجموعة تعليميه في كلية طب الاسنان في جامعه مشهد. لقد حصلنا على معدل CR في الاختبارات النهاية في الفصل الاول من العام الدراسي ٩٠-٩١. يساوي ٩٥.٢٪ و كانت معدلات CC و CV للاختبارات تعادل ٧٦.٨٧٪ و ٨٦.٠٤٪.

الإستنتاج: كانت النتيجه ايجابيه و لكن كان مستوى غطاء اهداف CCو CR في هذه الدراسه اقل من المتوقع في بعض المجموعات التعليمية و إن اجراء معامل تعليمية لتثقيف اعضاء الهيئه العلميه في مجال ادراك اهمية مستوى الدقة في صياغة الاختبارات يكون ضروري.

كلمات المفتاح: اختبار، دقة، دقة المحتوا، تناسب المحتوى، غطاء المحتوى، كلية طب الاسنان

تعيين روایی محتوای آزمون های کتبی برگزار شده در دانشکده دندانپزشکی مشهد در سال تحصیلی ۹۱-۹۰

مقدمه: روایی اصطلاحی است که به هدفی که آزمون برای تحقق بخشیدن به آن درست شده است اشاره می کند. آزمون های دانشگاهی ابزار مهمی جهت ارزیابی یادگیری و توانمندی دانشجویان هستند، لذا بررسی روایی و کفایت این آزمون ها ضروری به نظر می رسد. هدف این مطالعه تعیین روایی محتوای آزمون های کتبی برگزار شده در دانشکده دندانپزشکی مشهد در نیم سال اول سال تحصیلی ۹۱-۹۰ می باشد.

مواد و روش: در این مطالعه توصیفی و مقطعی که با همکاری دفتر توسعه آموزش (EDO) دانشکده دندانپزشکی مشهد انجام گردید، جهت بررسی روایی محتوای آزمون ها از کمی سازی روایی محتوا به روش مسیک استفاده شد. بدین منظور از هر گروه آموزشی یک آزمون نهایی برگزار شده در نیم سال اول سال تحصیلی ۹۱-۹۰ به صورت تصادفی انتخاب گردید. سپس جدول مشخصات (Blueprint) برای آن آزمون توسط یکی از اساتید گروه تهیه شد. در نهایت میزان تناسب محتوایی (CR)، پوشش محتوایی (CC) و روایی محتوا (CV) برای آن توسط اساتید طراح سوالات آزمون تعیین گردید.

نتایج: در این مطالعه ۱۱ گروه آموزشی دانشکده دندانپزشکی مشهد شرکت کردند. میانگین CR در آزمون های نهایی برگزار شده در نیمسال اول سال تحصیلی ۹۱-۹۰ برابر ۹۵/۲٪ بدست آمد. میانگین CC و CV برای آزمون ها به ترتیب برابر ۷۶/۸۷٪ و ۸۶/۰۴٪ بدست آمد.

نتیجه گیری: در این مطالعه میانگین CR و CC خوب ارزیابی شد اما میزان پوشش اهداف آموزشی در بعضی گروه های آموزشی کمتر از حد انتظار بود و برگزاری کارگاه های آموزشی جهت آشنایی هیات علمی با اهمیت روایی محتوایی آزمونها موجب طراحی دقیقتر و بهتر آزمونها خواهد شد

واژگان کلیدی: آزمون-روایی- روایی محتوا-تناسب محتوا-پوشش محتوا-دانشکده دندانپزشکی

مشہد ڈینٹل کالج میں دوہزار گیارہ میں منعقدہ کتبی ٹسٹ کے مضامین کی علمی حیثیت کا تعین کرنا

بیک گراؤنڈ: علمی حیثیت کا تعین اس وجہ سے کیا جاتا ہےکہ ٹسٹ کے اہداف پورے ہوں۔ یونیورسٹی ٹسٹ طلباء کی علمی کارکردگی کی جانچ کے لئے مفید وسیلہ ہیں لہذا ان کی علمی حیثیت کا تعین کیاجانا نہایت ضروری ہوتا ہے اس تحقیق کا ہدف مشہد ڈینٹل کالج میں ایجوکیشن ڈیولپمنٹ آرگنائزیشن کا تعاون حاصل تھا۔ ٹسٹوں کی علمی حیثیت کا تعین کرنے کے لئے کنٹنٹ کوریج اور کنٹنٹ ریلوینس کا تعین کیا گیا اس کے بعد کنٹنٹ ویلیڈیٹی کا میسیک طریقے سے حساب کیا گیا۔

روش: اس تحقیق میں مشہد ڈینٹل کالج میں ایجوکیشن ڈیولپمنٹ آرگنائزیشن کا تعاون حاصل تھا۔ ٹسٹوں کی علمی حیثیت کا تعین کرنے کے لئے کنٹنٹ کوریج اور کنٹنٹ ریلوینس کا تعین کیا گیا اس کے بعد کنٹنٹ ویلیڈیٹی کا میسیک طریقے سے حساب کیا گیا۔

نتیجہ: اس تحقیق میں ڈینٹل کالج کے گیارہ گروپوں نے شرکت کی تھی اوسط سی آر پہلے نصف سال میں پچانوے اعشاریہ دو تھا جبکہ اوسط سی سی اور سی وی ستاسی اعشاریہ سات چھے فیصد اور چھیاسی اعشاریہ صفر چار فیصد تھا۔

سفارش: اس تحقیق سے پتہ چلا ہےکہ اوسط سی آر اور سیسی ٹھیک رہا ہے جبکہ بعض گروپوں میں توقع سے کم رہا ہے جس کا تدارک کرنے کےلئے تعلیمی ورک شاپ لگائی جانی چاہیں۔ اس کے نتیجے میں ٹسٹ اور امتحان کے لئے بہتر سے بہتر سوالنامے بنائے جاسکتے ہیں۔

کلیدی الفاظ: ٹسٹ، علمی حیثیت، ڈینٹل کالج ۔

INTRODUCTION

Evaluation is a systematic determination of subjects' merit, value, and significance, which uses criteria governed by a set of standards. Evaluation can help an organization, program, or project to ascertain the degree of achievement or value in regard to its aims and objectives (1).

The main purpose of evaluation is determining the quality of a program to assess the effects of certain activities in three areas: knowledge, attitude, and behavior. Questionnaire, interview, and observation are three methods of collecting evaluation information (1).

Training evaluation is a continual and systematic process of assessing the worthiness and value of a training course, program, or activity. It is an essential objective of gathering quantitative and qualitative data regarding the effectiveness of training. Various components of training and also continuation, correction, or elimination of the training are determined by the obtained results of evaluation (2).

Training evaluation is the most important and integral part of training, if it is continuously tested and is based on scientific principles. In fact, the obtained results of evaluation should determine whether training assists the organization or program to improve its performance (3).

Today, the high prevalence of dental or oral diseases and socioeconomic complications greatly affect public health. Therefore, we need criteria to assess whether our training is valuable and applicable for the development and modification of dental knowledge and skills (4). There are various methods for training evaluation, and determining the validity and reliability of exam questions is one of the useful methods. The content validity indicates whether the exam questions represent the content and objectives of a specific topic; to wit, a written test has content validity if the questions represent all the objectives and course content (5). The evaluation of exam questions as a measurement tool for assessing students' learning is one of the main aspects of training evaluation. The examination questions should be carefully evaluated to assess the rate of students' learning (6).

Academic tests are useful tools for evaluating the knowledge and capabilities of students. Therefore, in order to be certain about the quality of questions, it is essential to assess the reliability and validity of these tests (7). This study aimed to assess the content validity of final exams in Mashhad Dental School. In addition, the practical purpose of this study was raising the quality of written exams of Mashhad Dental School.

METHODS

The present descriptive, cross-sectional study was conducted in order to determine the content validity of final exams in Mashhad Dental School, located in Northeast of Iran, in years 2011-2012. Inclusion criteria were: one of the multiple choice questions from following 11 departments: departments of pediatric dentistry, dental prosthesis, maxillofacial radiology, preiodontics, endodontics, orthodontics, restorative dentistry, maxillofacial surgery, oral and maxillofacial diseases, oral health and community dentistry, and pathology. Exclusion criteria were: lack of cooperation in educational groups of Mashhad Dental School. From all educational groups, one final exam was selected to determine the amount of the content and learning objectives coverage by the tests. For this purpose, one of the faculty members of each group was selected as the representative, and was provided with the required explanations to fill the related questionnaire. In this study, a six-item questionnaire was given to each of the deputy, and was completed by the group members and representatives.

The items included in these questionnaires were as follows:
1) Determine the percentage of items related to educational objectives.
2) Determine the percentage of educational objectives covered by at least one exam question.
3) What is the percentage of content validity of the exam?
4) How many items were unrelated to educational objectives?
5) In which domains of educational objectives, there were no questions associated with educational objectives?
6) According to the faculties, which domains of education are considered important (based on the number of exam questions)?
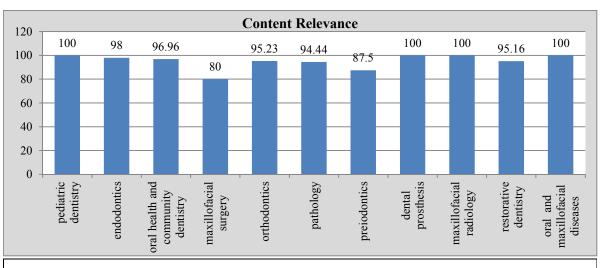

Chart 1. The percentage of content relevance of exams in different departments

The Messick method was used to assess the content validity of Dental School exams. In this study, the percentage of content coverage (CC), and the percentage of content relevance (CR) of each test item were determined by dental school departments. Afterwards, the percentage of content validity (CV) was calculated using Messick method; in fact, content validity is the mean of CC and CR.

In Messick method, a blueprint was used to determine the variables. In this study, the blueprint was assigned to the educational objectives and the syllabus content. Finally, the obtained results were presented to the department of dental school education to promote exam quality. We used frequency distribution table and diagrams in order to assess the content validity of final exams.

RESULTS

In total, 68 faculty members who were exam developers participated in the current study. The obtained results of each department were reported by the representatives.

According to the results of this study, the mean percentage of CR and CC was 95.2% and 76.87 in the final exams of all departments, respectively. In addition, the mean percentage of content validity was 86.04 in exams of different groups. The percentages of items related to the educational objectives, which is the percentage of CR of exams of different departments, are indicated in Chart 1.
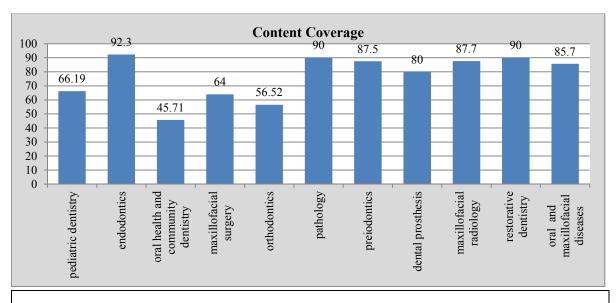

Chart 2. The percentage of content coverage of exams in different departments
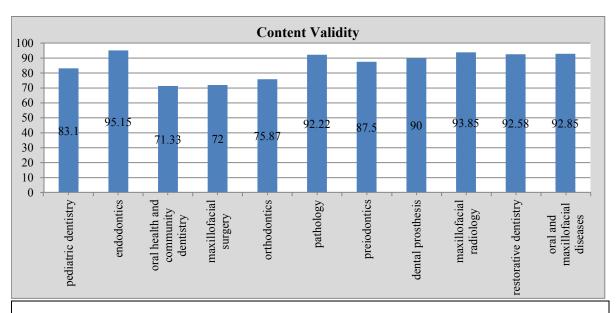

Chart 3. The percentage of the content validity of exams in different departments

The percentages of educational objectives covered by final exam items (the percentage of CC of exams) in different departments are presented in Chart 2. In addition, the CV percentages of exams in different departments are shown in Chart 3.

According to the results, one question item of the exams of both endodontics and oral health and community dentistry departments was unrelated to educational objectives. In addition, three items of both pathology and restorative dentistry departments were unrelated to educational objective. Two, four, and five items of the exams of orthodontics, maxillofacial surgery, and preiodontics departments were unrelated to the educational objectives, respectively. None of the items in the final exams of pediatric dentistry, dental prosthesis, maxillofacial radiology, and oral and maxillofacial diseases were unrelated to the objectives.

Finally, we determined the most important educational domains in different departments of Mashhad Dental School. The percentages of questions without CC of educational domains are demonstrated in Chart 4.

The radiographic and clinical diagnosis of complications related to dental trauma to primary teeth on the permanent dentition, selecting the type of spacer, treatment of luxation injuries, and preparation techniques for managing avulsed tooth were the most significant educational domains of the department of pediatric dentistry. The most important learning domains of the department of endodontics included root canal preparation techniques, and radiography in endodontic treatment and internal anatomy of the teeth. Oral public health, public health, socioeconomic consequences of tooth decay, and types of epidemiological studies were the most important educational domains of the department of oral health and community dentistry.

Also, the most significant educational domains of the department of maxillofacial surgery included flap surgery, pre-prosthetic surgery, and familiarity with dental equipments. The embryological origin of the formation of different facial components, developmental patterns of cranial base, complex nasolacrimal duct obstructions¸ and mandible, and the mechanism of tooth development (before and after the development in the mouth) were the most important educational domains of the department of orthodontics.

The important educational domains of the department of pathology consisted of chronic mucocutaneous candidiasis, developmental disorders, and microbial diseases. In addition, the gingiva, plaque and calculus, and the etiology of periodontal diseases were the important educational domains of the department of preiodontics.

The important educational domains of the department of dental prosthesis included retention in partial dentures and the related factors, forms of indirect retainers and the related factors, and components of partial dentures. The most important educational domains of the department of maxillofacial radiology were knowledge of the factors affecting image quality and film processing, and factors affecting X-ray and tissue sensitivity to radiation. Also, complex amalgam restorations, examination, diagnosis and treatment planning, and pulp-protection materials were the most important educational domains of the department of restorative dentistry. Finally, laboratory tests, red and white lesions, and exophytic lesions were the most important educational domains of the department of oral and maxillofacial diseases.
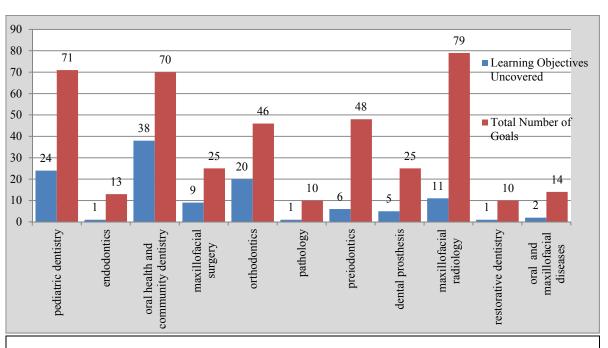

Chart 4. Educational domains without content coverage

DISCUSSION

The results of this study showed that the mean percentage of CR and CC of final exams was 95.2% and 76.87% in all educational groups, respectively. Also, the mean of CV of final exams was 86.04% in all groups.

The results of the current study were consistent with the study of Kazemi and Ehsanpour, which assessed the content validity of exam questions of theoretical courses of midwifery. In addition, the degree of compliance between exam questions and the level of students' learning was assessed in the mentioned study. The results indicated that the content validity and face validity of exam questions were 92.38% and 84.45%, respectively. However, the level of students' learning was not consistent with the expected level (10).

Also, the results of the present study were in consistence with studies of Edhami et al. (11), Heravi et al. (12), and Hasnzadeh et al. (13).

Jahani Shorab in his study aimed to assess the reliability and validity of assessment tools of the final examination of the department of nursing and midwifery. The results of that study indicated that the assessment tools of comprehensive examination were reliable tools for evaluating students at the time of graduation (14).

The obtained results of this study indicated that there was a large difference between the percentage of CR and CC. The percentage of CR ranged between 80% and 100%, though the percentage of CC ranged between 45.71% and 92.3%. Therefore, CC plays a more significant role in comparison with CR in determining the percentage of content validity.

Also, other results of this study showed that there was no relationship between the number of faculty members and number of educational objectives. Also, there was no relationship between the gender of faculty members and the number of educational objectives.

Consistence with our study, Kazemi and Ehsanpour (10) and Jandaghi (15) used the Messick method to assess content validity. However, Hsanzade et al. (13) and Jahani Shorab et al. (14) used the Lawshe method for their assessment of content validity.

Najar and Abedi in their study assessed the rate of awareness of faculty members regarding measurement methods of reliability and validity of exam questions. The results indicated that 80% of faculty members had sufficient knowledge about content validity to determine the validity of exam questions. However, a small percentage of subjects had sufficient knowledge about the methods of assessing reliability and validity of exam questions (16).

Shaban and Ramezani Badr evaluated the content validity of exam questions of nursing and midwifery of Tehran University of Medical Sciences. According to their study, the educational pamphlets to faculty members have a considerable effect on improving the designed questions (17).

This study concluded that the mean percentage of CR and CC of final exams was good (95.2% and 76.87%) in all educational groups. Regarding the validity of the exams, evaluating the content aspect and the supervision of the standard designation of questions by the EDO could lead to an improvement in the quality of the examinations and their validity, which was suggested to be generalized for all other programmed exams. Moreover, teachers' awareness of the importance of content validity and calculation methods are very effective in designing efficient exams with greater sensitivity.

ACKNOWLEDGEMENTS

REFERENCES

1. Afshari R, Sarabadani J, Ghazvini K, mousavi R, Karimi mouneghi H,Hashemi S E, etal. Student educational evaluation in medical sciences. 1st ed. Mashhad, Mashhad University of Medical Sciences and Health Services. 2012:12-44. (Persian)

2. Jarahi L, Mousavi Bazaz SM, Erfanian MR,Khajedalouee M. Evaluation of Medical Students during Internship Based on Performance Rating Tables. FMEJ, 2015;3(5):12-15

3. Alvarez K, Salas E, Garofano CM. An integrated model of training evaluation and effectiveness. Human Resource Development Review December 2004; 3: 385-416.

4. Eseryel D. Approaches to Evaluation of Training: Theory & Practice. Educational Technology & Society, 2002; 5(2): ISSN 1436-4522.

5. Mohebbi SZ, Esfandiyari M. the effect of the theoretical course of community oral health on the oral health prevention knowledge, attitude and self-reported practice in dental school students at Tehran University of Medical Sciences. Journal of Dentistry of Tehran University of Medical Sciences 2011; 24(4): 245-25. (Persian)

6. Seif AA. Educational measurement, assessment, and evaluation. 6thed. Tehran, Iran: Doran publication; 2012, 30-40, 529-545. (Persian)

7. Polit DF, Beck CT.The content validity index: Are you sure you know what's being reported? Critique and recommendations. Research in Nursing & Health, 2006; 29 (5): 489-497.

8. Mahmoodi MR,Baneshi MR,Alizadeh SM. Influence of Assessment Method Selection in Studying and Learning Approaches: Is It Necessary to Change Assessment Style?. 2014;4(2):35-40

9. Sireci S, Faulkner-Bond M. Validity evidence based on test content. Psicothema. 2014; 26(1):100-7.

10. Kazemi A,Ehsanpour S. The analysis of examination questions of theoretical courses in Esfahan University of medical sciences. Iranian Journal of Education at Medical Sciences, 2010; 10(5): 643-650. (Persian)

11. Edhami A, Nakhaee N. Fasihi Harandi T, Fatahi Z. Preliminary Study of the validity and relability of faculty evaluation forms (evaluation of teaching by students) in Kerman University of medical sciences in 2002-2003. Journal of medical education and development. 2005; 2:121-129. (Persian)

12. Heravi Karimoy M, Anooshe M, Forughan M, Sheykhi M, Hajizade E. The psychometric properties and design of Elder Abuse Questionnaire. Iranian Journal of Ageing. 2010; 15 (5): 7-22. (Persian)

13. Hasanzade Rangi N, allahyari T, Khosravi Y, Zaeri F, Saremi M. Designing of questionnaire of The cognitive failures questionnaire: Determining the validity and reliability of the tool. Iran occupational health journal. 2012; 8(4): 29-40. (Persian)

14. Jahani Shorab N, Golmakani N, Mazloom SR, Mirzakhani K, Azhari S, Navaeeyan A. The validity and reliability of evaluation tools of end of course examination of students of nursing and midwifery of Mashhad univerity. Iranian journal of education at medical sciences, 2011; 11(5): 485-495. (Persian)

15. Jandaghi G. assessment of validity, relabiliy and difficalty indices for teacher-built physis exam questions in first year high school. Arts and social science j ournal. 2011; 5(11): 651-654.

16. Najar SH, Abedi P. knowledge and practic of faculty members in assessing the validity and reliaility of examinations. J Ghazvin University of Medical Siences. 2004; 30(1): 46-50. (Persian)

17. Shaban M, Ramezanibadr F. the Impact analysis of the final examinations on the quality of design of questions. Journal of Nursing and Midwifery of Tehran University of Medical Sciences. 2007; 13(1): 5-15. (Persian)