Development Strategies in Medical Education

Summr 2024. Vol 11. Issue 2

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

Research Article

Comparing the extended response matching test with the Four-Choice Test: Results of a Quasi-Experimental Study

Shima Sadat ZarifNahad[1] , * Majid Khadem Rezaiyan[2&3] , Javad Sarabadani[4] , Mina AkbariRad[5] , Mahnaz Boroumand Rezazadeh[6]

1. Department of Medical Education, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.
2. Medical Sciences Education Research Center, Mashhad University of Medical Sciences, Mashhad, Iran.
3. Department of Community Medicine, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.
4. Department of oral and maxillofacial diseases, School of Dentistry, Mashhad University of Medical Sciences, Mashhad, Iran.
5. Department of Internal Diseases, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.
6. Department of Obstetrics and Gynecology, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.

Use your device to scan and read the article online

Citation: ZarifNahad SS, Khadem Rezaiyan M, Sarabadani J, AkbariRad M, Boroumand Rezazadeh M. Comparing the extended response matching test with the Four-Choice Test: Results of a Quasi-Experimental Study. Development Strategies in Medical Education. 2024; 11(2):151-164. [In Persian]

Article Info:
Received: 11 Aug 2024
Accepted: 10 Sep 2024
Available Online: 20 Sep 2024

ABSTRACT

Introduction: Collective assessment methods in the medical field primarily focus on multiple-choice questions (MCQs) designed to measure the recall of information and basic facts. However, extended matching questions (EMQs), as an advanced type of multiple-choice test, have the capability to assess higher cognitive levels such as problem-solving and the application of knowledge. In this study, aimed at comparing the effectiveness and impact of EMQs and MCQs on the performance and satisfaction of medical students, these two types of assessments were evaluated among students at Mashhad University of Medical Sciences. The results of this study could significantly contribute to improving the quality of assessment in medical education and identifying more effective methods for measuring the clinical abilities of students.

Methods: This research was conducted as a semi-experimental intervention in 2021 on obstetrics and gynecology and internal medicine students at Mashhad University of Medical Sciences. The sample for this study was obtained through convenience sampling, including students in residency, internship, and clerkship programs. A multiple-choice questionnaire (MCQ) was used in the control group, while an extended matching question (EMQ) format was employed in the intervention group. The validity of the test questions was confirmed by a panel consisting of five subject matter experts and one medical education specialist. The collected data were analyzed using SPSS software version 25. To compare the effectiveness of the EMQ and MCQ tests in distinguishing between students at three different educational levels, an analysis of variance (ANOVA) was conducted.

Results: The A total of 176 individuals participated, with 87 in the intervention group and 89 in the control group. The baseline variables showed no statistically significant differences between the two groups. The EMQ score (17.1±2.5) was higher than the MCQ score

*** Corresponding Author:**
Dr Majid Khadem Rezaiyan
Address: Mashhad University of Medical Sciences, Mashhad, Iran.
Tel: +98 5138829262
E-mail: khademrm@mums.ac.ir

Copyright © 2024 The Author[s];
This is an open access article distributed under the terms of the Creative Commons Attribution License [CC-By-NC: https://creativecommons.org/licenses/by-nc/4.0/legalcode.en], which permits use, distribution, and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Development Strategies in Medical Education

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

ABSTRACT

(16.6±2.6), although this difference was not statistically significant. Additionally, student satisfaction with EMQ was greater than with MCQ (17.5 ±2.5 compared to 16.2±3.3). The EMQ test score had a significant inverse correlation with the entrance exam score (r = -0.399, p = 0.011), while this correlation was not observed for the MCQ test (r = -0.258, p = 0.095).

Discussion: According to learning and educational psychology, monitoring and evaluation are considered the final link in educational activities and complement the teaching process. It is noteworthy that the behavior of some medical students indicates that the method of evaluation can stimulate learning. Therefore, efforts to find the best evaluation method continue. Currently, Multiple Choice Questions (MCQs) are significantly used in medical universities due to their ability to assess a large number of students and evaluate a wide range of educational content with high objectivity. However, most specialists do not recommend this method. The results indicate that student performance and satisfaction in Extended Matching Questions (EMQs) are higher compared to MCQs, which may be due to a more accurate assessment of students and a more objective evaluation of educational content in the extensive matching test.

Key Words:

Educational Measurement, Medical Education, Medical Students, Medicine.

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

Extended Abstract

Introduction:

In learning and educational psychology, monitoring and evaluation are recognized as integral components of the educational process, completing the teaching cycle. Accurate assessment of core competencies, such as medical knowledge, is essential for providing reliable feedback and driving improvements in medical education. Multiple-choice questions (MCQs) are a commonly used assessment tool in medical education, though they have been associated with promoting surface learning strategies. While MCQs are widely employed in medical universities, particularly in large classes, their efficacy is debated. Many specialists argue that MCQs primarily assess recall and fail to evaluate higher-order cognitive processes like reasoning and problem-solving. Extended matching questions (EMQs) are a recommended alternative for assessing clinical judgment and reasoning. A specialized form of MCQ, EMQs are designed to evaluate higher cognitive levels, including problem-solving. Existing research supports the validity and reliability of EMQs, suggesting that they provide a more sensitive measure of student progress compared to MCQs and effectively assess both factual knowledge and clinical reasoning. The format of EMQs, with numerous plausible options related to the question topic, reduces the likelihood of guessing, particularly for weaker students. Despite these advantages, a review of the literature suggests that no prior study has directly compared EMQs and MCQs among medical students in Iran. Therefore, this study aims to compare the effectiveness of EMQs and MCQs in assessing medical students at Mashhad University of Medical Sciences, examining both student performance and satisfaction with each test format.

Methods:

This This quasi-experimental, interventional study was conducted at Mashhad University of Medical Sciences in 2021, in collaboration with the Departments of Obstetrics and Gynecology and Internal Medicine. The study population comprised residents (second year and above), medical interns (at least six months of departmental experience), and physiopathology/clerkship students (at least three months in their respective programs). The exclusion criterion was unwillingness to participate. Participants were assigned to either an intervention or control group. Residents were randomized by hospital (Ghaem (AJ) and Imam Reza (AS)), while medical students were randomized during two departmental entry periods. The intervention group completed an EMQ test, and the control group completed an MCQ test. The MCQ test, comprising 20 questions, was designed to match the EMQ test in terms of question number, difficulty, and conceptual content. MCQs were primarily designed using the “best answer” format. Five specialists from the relevant fields and one medical education specialist reviewed the questions. A six-member expert panel ensured consistency between the two tests regarding difficulty, content, and topic coverage. Questions in both tests were categorized as easy, medium, or difficult. Neither test employed negative marking, and both had a 90-minute time limit. Data were analyzed using SPSS version 25. Quantitative variables are presented as mean, standard deviation, and range; qualitative variables as frequency and percentage. The Student’s t-test was used to compare quantitative variables between groups. ANOVA was employed to assess the discriminative power of the EMQ and MCQ tests across the three educational levels. All tests were two-tailed, with a significance level of $p < 0.05$.

Results:

A total of 176 students participated in the study (87 in the EMQ group and 89 in the MCQ group). No statistically significant differences were found between the two groups regarding educational level (internship, residency, and assistantship), gender, quota type, age, entrance exam rank, or time elapsed since the start of professional doctoral studies. Student satisfaction with the EMQ exam was significantly higher than with the MCQ exam (mean ± SD: 17.5 ± 2.5 vs. 16.2 ± 3.3, $p = 0.004$). While satisfaction with the EMQ exam was greater across all three educational levels (Figure 2), this difference reached statistical significance only at the assistantship level ($p = 0.040$). Finally, EMQ exam scores demonstrated a significant inverse correlation with entrance exam rank ($r = -0.399$, $p = 0.011$), whereas no such correlation was observed for MCQ exam scores ($r = -0.258$, $p = 0.095$).

Development Strategies in Medical Education

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

Conclusion:

The Final evaluation methods in medical education frequently rely on multiple-choice tests, which primarily assess the recall of discrete facts. However, a more holistic evaluation of medical students is needed, particularly one that assesses their ability to apply knowledge in clinical settings. A review of the medical education assessment literature suggests that Extended Matching Questions (EMQs) have demonstrated greater effectiveness in evaluating medical students, with high levels of student satisfaction. While numerous studies have compared teaching methodologies, research on evaluation methods is comparatively limited, often consisting of reviews rather than direct comparisons of assessment tools. This study addressed this gap in the literature. Based on the findings, it is recommended that EMQ tests be implemented for student assessment in clinical environments due to their capacity to evaluate higher-order cognitive skills and the high levels of student satisfaction associated with this format. A limitation of this study is the recruitment of specialty residents from only two educational groups. While these are major groups, broadening the diversity of educational groups could enhance the generalizability of the findings. Conversely, the inclusion of three distinct educational levels (internship, residency, and specialty training) represents a strength, providing a more comprehensive understanding of the impact of the two methods across the educational spectrum. Future research comparing these two assessment methods in other medical specialties is recommended.

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

مقاله پژوهشی

مقایسه آزمون جورکردنی گسترده پاسخ با آزمون چهارگزینه‌ای: نتایج یک مطالعه نیمه تجربی

شیماسادات ظریف‌نهاد[۱]، * مجید خادم‌رضائیان[۲و۳]، جواد سرآبادانی[٤]، مینا اکبری‌راد[٥]، مهناز برومند رضازاده[٦]

۱. گروه آموزش پزشکی، دانشکده پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران.
۲. مرکز تحقیقات آموزش علوم پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران.
۳. گروه پزشکی اجتماعی، دانشکده پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران.
٤. گروه بیماری‌های دهان و تشخیص، دانشکده دندانپزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران.
٥. گروه بیماری‌های داخلی، دانشکده پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران.
٦. گروه زنان و زایمان، دانشکده پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران.

Use your device to scan and read the article online

Citation: ZarifNahad SS, Khadem Rezaiyan M, Sarabadani J, AkbariRad M, Boroumand Rezazadeh M. Comparing the extended response matching test with the Four-Choice Test: Results of a Quasi-Experimental Study. Development Strategies in Medical Education. 2024; 11(2):151-164. [In Persian]

اطلاعات مقاله

تاریخ دریافت: ۲۱ مرداد ۱۴۰۳

تاریخ پذیرش: ۲۰ شهریور ۱۴۰۳

تاریخ انتشار: ۳۰ شهریور ۱۴۰۳

چکیده

هدف: روش‌های ارزشیابی جمعی در حوزه پزشکی، عمدتاً بر آزمون‌های چندگزینه‌ای (MCQ) متمرکز هستند که به منظور سنجش یادآوری اطلاعات و حقایق پایه طراحی شده‌اند. با این حال، آزمون‌های جورکردنی گسترده پاسخ (EMQ) به عنوان نوعی آزمون چندگزینه‌ای پیشرفته، توانایی ارزیابی سطوح بالاتر شناختی مانند حل مسئله و کاربرد دانش را دارا هستند. در این پژوهش، با هدف مقایسه کارایی و تأثیر آزمون‌های EMQ و MCQ بر عملکرد و رضایت دانشجویان پزشکی، این دو نوع آزمون در دانشجویان دانشگاه علوم پزشکی مشهد مورد ارزیابی قرار گرفت. نتایج این مطالعه می‌تواند به بهبود کیفیت ارزشیابی در آموزش پزشکی و شناسایی روش‌های مؤثرتر برای سنجش توانایی‌های بالینی دانشجویان کمک شایانی نماید.

روش‌ها: این پژوهش مداخله‌ای نیمه‌تجربی در سال ۱۴۰۰ بر روی دانشجویان رشته‌های زنان و زایمان و بیماری‌های داخلی دانشگاه علوم پزشکی مشهد انجام شد. داده‌ها از طریق نمونه‌گیری در دسترس، (دانشجویان دستیاری، کارورزی و کارآموزی) انجام گردید. در گروه کنترل، از آزمون چهارگزینه‌ای و در گروه مداخله، از آزمون جورکردنی گسترده پاسخ استفاده شد. روایی سوالات آزمون توسط پنلی متشکل از ۵ متخصص رشته‌ای و یک متخصص آموزش پزشکی تأیید گردید. داده‌های با استفاده از نرم‌افزار SPSS تحلیل شدند. به منظور مقایسه توانایی آزمون‌های EMQ و MCQ در تمایز بین دانشجویان سه مقطع تحصیلی، از آزمون تحلیل واریانس (ANOVA) استفاده شد.

یافته‌ها: میانگین در مجموع ۱۷۶ نفر مشارکت نمودند که ۸۷ نفر در گروه مداخله و ۸۹ نفر در گروه کنترل بودند. متغیرهای پایه بین دو گروه تفاوت آماری معناداری نداشت. نمره EMQ (۲/۵±۱۷/۱) بیشتر از MCQ (۲/۶±۱۶/۶) بود هرچند این تفاوت از نظر آماری معنی دار نبود. همچنین رضایت دانشجویان از EMQ بیشتر از MCQ بود (۲/۵±۱۷/۵ در مقابل ۳/۳±۱۶/۲). نمره آزمون EMQ با رتبه کنکور همبستگی معکوس و معنی‌داری داشت (r= -۰/۳۹۹, P=۰/۰۱۱) اما این همبستگی برای آزمون MCQ مشاهده نشد (r= -۰/۲۵۸, P=۰/۰۹۵).

*نویسنده مسئول:
دکتر مجید خادم رضائیان
نشانی: دانشگاه علوم پزشکی مشهد، مشهد، ایران.
تلفن: 5138829262 98+
پست الکترونیک: khademrm@mums.ac.ir

Copyright © 2024 The Author[s];
This is an open access article distributed under the terms of the Creative Commons Attribution License [CC-By-NC: https://creativecommons.org/licenses/by-nc/4.0/legalcode.en], which permits use, distribution, and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

چکیده

نتیجه‌گیری: براساس روانشناسی یادگیری و آموزشی، نظارت و ارزشیابی حلقه آخر فعالیت‌های آموزشی و تکمیل کننده انجام آموزش محسوب می‌شود نکته قابل تامل آن است که رفتار برخی دانشجویان پزشکی حاکی از آن است که نحوه ارزیابی می‌تواند منجر به تحریک یادگیری شود بر همین اساس تلاش برای یافتن بهترین روش ارزیابی همچنان ادامه دارد. در حال حاضر MCQ به دلیل توانایی ارزیابی تعداد زیادی دانشجو و همچنین سنجش طیف وسیعی از محتوای آموزشی با عینیت بالا، به طور چشمگیری در دانشگاه‌های علوم پزشکی استفاده می‌شوند با این حال بیشتر متخصصین استفاده از این روش را توصیه نمی‌کنند که با توجه به نتایج به دست آمده مشخص گردید که عملکرد و رضایت دانشجویان در EMQ نسبت MCQ بیشتر می‌باشد که می‌تواند به دلیل ارزیابی دقیق‌تر دانشجویان و سنجش عینی‌تر محتوای آموزشی در آزمون جورکردنی گسترده پاسخ باشد.

کلیدواژه‌ها :
اندازه گیری آموزشی، آموزش پزشکی، دانشجو پزشکی، پزشکی.

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

مقدمه:

براساس روانشناسی یادگیری و آموزشی، نظارت و ارزشیابی حلقه آخر فعالیت‌های آموزشی و تکمیل کننده انجام آموزش محسوب می‌شود. اندازه‌گیری کافی شایستگی‌های اصلی مانند دانش پزشکی یک جزء ضروری برای ارزیابی، ارائه بازخورد قابل اعتماد و بهبود آموزش پزشکی است. روش‌های ارزیابی جمع‌بندی که برای سنجش دانش پزشکی استفاده می‌شود، عمدتاً شامل سؤالات چند گزینه‌ای (MCQ) که برای ارزیابی یادآوری دانش یا حقایق پزشکی است. این امر نیاز به ارزشیابی دقیق‌تر دانشجویان پزشکی را با اندازه‌گیری نتایجی که منعکس‌کننده فرآیندهای مرتبه بالاتر مانند کاربرد دانش است، برجسته می‌کند [۱].

در کل می‌توان گفت که از این میان رایج‌ترین ابزار سنجش و ارزیابی، آزمون‌های MCQ می‌باشد که به طور گسترده در آموزش پزشکی استفاده می شوند و می توانند استراتژی‌های یادگیری سطحی را ارتقا دهند [۲]. سوالات چند گزینه‌ای خود به سه دسته تقسیم می‌شوند: ۱- نوع تنها یک گزینه درست: دراین نوع یک گزینه کاملا درست و بقیه کاملا غلط هستند. ۲- نوع بهترین گزینه: همه گزینه‌ها ممکن است درست باشند اما یک گزینه صحیح‌ترین جواب است. ۳- نوع منفی: تنه سوال به صورت منفی بیان می‌شود و همه گزینه‌ها به جز یکی غلط هستند. یکی از مشکلات تهیه سوال‌های چند گزینه‌ای این است که تهیه تعدادی گزینه انحرافی غلط که به ظاهر درست جلوه کند عمدتا مشکل است [۳].

آزمون جورکردنی گسترده پاسخ (EMQ) نشان دهنده نوع خاصی از آزمون‌های MCQ می‌باشد که برای ارزیابی سطوح بالاتر شناختی مانند حل مسئله طراحی شده است [۴]. علاوه بر آن عمدتا سوالات در حد سنجش اطلاعات (یادآوری) طراحی می‌شوند. از سوی دیگر در ارزیابی مهارت های حل مسئله دانشجویان قدرت کمتری دارند. اما از مزایای آزمون‌ها چند گزینه‌ای آن است که در یک زمان می‌توان تعداد زیادی سوال را مورد استفاده قرار داد. بنابراین، امکان ارزیابی نمونه نسبتا بزرگی از محتوا درس و هدف‌های آموزشی را فراهم می‌آورد. با این وجود ممکن است یافتن مجموعه‌ای از پرسش‌ها و مجموعه‌ای از پاسخ‌های همگون مشکل باشد [۵].

نکته قابل تامل آن است که رفتار برخی دانشجویان پزشکی حاکی از آن است که نحوه ارزیابی می‌تواند منجر به تحریک یادگیری شود [۶]. بر همین اساس تلاش برای یافتن بهترین روش ارزیابی همچنان ادامه دارد. هرچند به جرأت می توان گفت تنها با تجمیع تمام روش‌های ارزیابی می‌توان به مطلوب ترین نتیجه رسید. در حال حاضر آزمون‌های چهارگزینه‌ای به طور چشم‌گیری در دانشگاه‌های علوم پزشکی که عمدتا کلاس‌ها دارای جمعیت زیاد می‌باشد، استفاده می‌شوند [۷]. روش های مختلف برای تعیین نقاط برش مناسب در آزمون های پیشرفت تحصیلی نیز مورد بحث قرار گرفته است [۸]. با این حال بیشتر متخصصین استفاده از این روش را توصیه نمی‌کنند. این متخصصین معتقدند که آزمون‌های چهارگزینه‌ای تنها یادآوری را سنجیده و فرآیندهای شناختی عالی‌تری مانند استدلال یا حل مساله را ارزیابی نمی‌کنند [۹]. یکی از بهترین ابزارهای توصیه شده برای ارزیابی قضاوت و استدلال بالینی آزمون‌های EMQ می‌باشند [۱۰]. این آزمون از چهار قسمت موضوع (تم)، سوال، پاسخ و راهنما، تشکیل شده است (۶). در مقالات مختلف مزایای زیر را برای این روش ارزیابی برشمرده‌اند: ۱- قالب تم‌ها به سازمان بندی آزمون کمک می‌کند. بدیهی است که استفاده از طرح درس برای فرآیند طراحی آزمون نیز کمک کننده خواهد بود. ۲- از آنجایی که سوالات در قالب تم‌ها یا موضوع کلی نگارش می‌شوند، مدرس می‌تواند سوالات زیاد را برای آن تم طراحی نموده و در نهایت برای برگزاری دفعات متعدد یک آزمون، از میان آنها به تصادف تعدادی سوال را برگزیند. ۳- سوالات خوب این امکان را فراهم می‌آورند تا کاربرد دانش سنجیده شود و نه فقط یادآوری صرف دروس. ۴- رویکرد نگارش این سوالات نظام‌مند است. این مساله به خصوص زمانی که افراد متعددی برای یک آزمون سوال طرح می‌کنند باید مدنظر قرار گیرد. ۵- لیست گسترده گزینه‌ها این امکان را فراهم می‌کند که تمام گزینه‌های مرتبط وارد آزمون شود.

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

این مساله منجر به کاهش حدس گزینه صحیح (مانند آنچه در سوالات چهارگزینه‌ای به فراوانی رخ می‌دهد) توسط دانشجوها خواهد شد. ۶- در نهایت نشان داد شده است که آزمون های مبتنی بر EMQ قدرت تمیز بالاتری نسبت به نسخه‌های ۲ یا ۵ گزینه‌ای همان سوالات دارند. در نتیجه توزیع نمرات گسترده‌تر بوده و پایایی بالاتری وجود خواهد داشت [۴،۶،۱۱،۱۲].

مطالعات اخیر از روایی و پایایی قابل قبول این روش حمایت می‌کنند [۶،۱۱،۱۲]؛ همچنین نشان می دهند که آزمون‌های مبتنی بر EMQ نسبت به آزمون‌های MCQ بهتر می‌توانند روند پیشرفت دانشجو را حین تحصیل پایش کنند، همچنین به خوبی می‌تواند دانش صرف و استدلال بالینی را ارزیابی کنند [۱۷-۱۳]. در این نوع آزمون به دلیل داشتن گزینه‌های زیاد و اینکه تمام گزینه‌ها مرتبط با موضوع سوال هستند، احتمال حدس زدن پاسخ صحیح (به خصوص برای دانشجویان ضعیف‌تر) به شدت کاهش می‌یابد. زمانی که از گزینه‌های زیادی استفاده شود، آزمون‌های EMQ شبیه آزمون‌های با پاسخ گسترده می‌شوند که دانشجو را مجبور به ساختن (ایجاد) پاسخ می‌کند تا اینکه تک تک گزینه‌ها را با سوال مورد نظر چک کرده و پاسخ صحیح را شناسایی کنند [۱۸]. با این وجود، برخی مطالعات نیز یافته‌های متناقضی را گزارش نموده‌اند. به طور مثال مطالعه تیجس و همکاران نشان داده است که آزمون چهارگزینه ای نسبت به آزمون EMQ توانایی بالاتری برای شناسایی دانشجویان ضعیف دارد. از طرف دیگر، آزمون EMQ نسبت به آزمون چهارگزینه ای توانایی بالاتری برای شناسایی دانشجویان قوی دارد [۷].

علی رغم مزایای متعدد ذکر شده در بالا بر اساس جستجوی محققین تا کنون، مطالعه‌ای به مقایسه این دو روش در دانشجویان پزشکی در ایران نپرداخته است، لذا در این مطالعه آزمون جورکردنی گسترده پاسخ با آزمون چهارگزینه‌ای در دانشجویان دانشگاه علوم پزشکی مشهد مقایسه می‌گردد تا تاثیر هر دو روش و همچنین نحوه‌ی عملکرد و رضایت دانشجویان در هر دو نوع آزمون سنجیده شود.

مواد و روش‌ها:

پژوهش این مطالعه مداخله‌ای (نیمه تجربی) در دانشگاه علوم پزشکی مشهد انجام شده است. معیار ورود عبارت بود از دستیاران سال ۲ و بالاتر، کارورزان پزشکی با سابقه حداقل ۶ ماه از ورود به بخش، دانشجویان مقطع فیزیوپاتولوژی یا کارآموزی با سابقه حداقل ۳ ماه تحصیل در مقطع مربوطه. معیار خروج عبارت بود از عدم تمایل به شرکت در مطالعه. تخصیص شرکت‌کنندگان به دو گروه به این صورت بود: دستیاران به صورت تصادفی بر اساس بیمارستان (قائم (عج) و امام رضا (ع)) و دانشجویان پزشکی عمومی به صورت تصادفی در دو دوره ورودی به بخش مربوطه به یکی از گروه‌های مداخله و کنترل اختصاص یافتند. در گروه مداخله از آزمون EMQ و در گروه کنترل از آزمون MCQ استفاده گردید. .

این طرح با همکاری گروهای زنان و زایمان و بیماری‌های داخلی که دانشجویانی از هر سه مقطع ذکر شده را داشتند، انجام شد. آزمون EMQ به صورت ۴ تم (هر یک حاوی ۵ سوال) که هر تم بین ۹ تا ۱۲ گزینه دارد، طراحی گردید. تم‌ها و سوالات توسط ۵ متخصص از همان رشته و یک متخصص آموزش پزشکی بررسی شد. هر سناریو مبتنی بر وضعیت بالینی و سوالات به منظور انتخاب بهترین اقدام تشخیصی یا درمانی در مدیریت بیمار مربوطه می‌باشد. هر گزینه ممکن است برای بیش از یک سوال صحیح باشد یا اصلا برای هیچ سوالی صحیح نباشد. آزمون MCQ نیز با تعداد سوالات مشابه (۲۰ سوال) و با سختی مشابه مبتنی بر مفاهیم مورد ارزیابی در آزمون EMQ طراحی و تا حد امکان سوالات آزمون چهارگزینه‌ای بر اساس روش بهترین پاسخ صحیح طراحی گردید. تیم ۶ نفره پانل خبرگان بر همسانی دو آزمون از لحاظ دشواری و محتوا و مباحثی که سوالات از آن طراحی می‌شود، نظارت داشت. در نهایت سوالات هر دو آزمون در سه سطح آسان، متوسط و دشوار دسته بندی گردید. هیچ یک از دو آزمون نمره منفی نداشت و مدت آزمون ۹۰ دقیقه بود. با توجه به اینکه شرکت‌کنندگان در طرح می‌دانستند که این آزمون به صورت آزمایشی برگزار می‌شود و لذا ممکن

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

بود در پاسخ‌دهی آنان سوگیری ایجاد کند ارزیابی گروه فیزیوپاتولوژی/ کارآموزان و کارورزان در یکی از امتحانات پایان بخش اجرا گردید. برای رعایت حقوق دانشجویان نمره نهایی دانشجویان از استاندارد آزمون تعیین شده است.

استاندارد نمره با روش جدید با معدل نمرات در ترم‌های گذشته مقایسه و در صورت تفاوت معنی‌دار برای جلوگیری از تضییع حق در معدل کل، با نظر گروه آموزشی مربوطه نمره بر اساس میانگین سنوات قبل نرمالایز و استاندارد گردید.

متغیرهای مورد سنجش، سن، جنس، مقطع تحصیلی، رتبه کنکور، سهمیه قبولی، تعداد ماه‌های سپری شده از شروع تحصیل، نمره آزمون، مدت زمان انجام آزمون، نمره رضایت از سطح آزمون و گزینه‌های انتخابی بود. قبل از برگزاری آزمون اصلی، یک جلسه توجیحی برای شرکت‌کنندگان برگزار شد تا با این روش آشنا شوند. حجم نمونه این مطالعه با استناد به یافته‌های مطالعه نظیم و همکاران در سال ۲۰۱۹ و با استفاده از فرمول مقایسه یک صفت کمی در دو جامعه محاسبه شد [۱۹]. با در نظر گرفتن خطای آلفای ۵ درصد و توان ۸۰ درصد، با در نظر گرفتن مقایسه نمره آزمون EMQ در دو گروه دستیاران و کارورزان پزشکی عمومی حجم نمونه معادل ۹ نفر در هر گروه و با در نظر گرفتن مقایسه نمره آزمون EMQ در دو گروه دستیاران و دانشجویان پزشکی عمومی حجم نمونه معادل ۱۸ نفر در هر گروه محاسبه گردید. بر همین اساس مقدار بزرگتر مدنظر قرار گرفته و با احتساب ۱۰ درصد ریزش، حداقل ۲۰ نفر در هر مقطع و هر گروه مطالعه (۶۰ نفر گروه مداخله و ۶۰ نفر گروه کنترل، در مجموع حداقل ۱۲۰ نفر) مورد ارزیابی قرار گرفتند. .

داده‌ها در نرم افزار SPSS نسخه ۲۵ وارد شد. متغیرهای کمی بوسیله میانگین، انحراف معیار، و دامنه و متغیرهای کیفی توسط فراوانی و درصد فراوانی گزارش گردیده است. مقایسه متغیرهای کمی بین دو گروه بوسیله آزمون تی دانشجویی سنجیده شده است. از آزمون ANOVA برای قدرت تمایز آزمون‌های EMQ و MCQ بین شرکت‌کنندگان در سه مقطع تحصیلی مورد ارزیابی استفاده گردید و تمامی آزمون‌ها دو طرفه و سطح معنی داری کمتر از ۰/۰۵ در نظر گرفته شده است.

یافته‌ها:

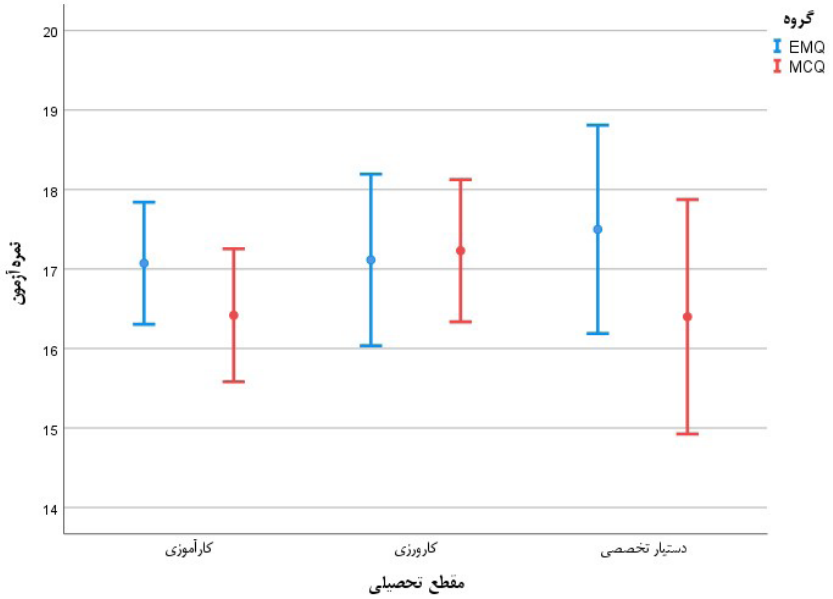
در مجموع ۱۷۶ دانشجو در این مطالعه شرکت کردند که ۸۷ نفر در گروه EMQ و ۸۹ نفر در گروه MCQ بودند. مقطع تحصیلی (کارآموزی، کارورزی و دستیاری)، جنسیت، نوع سهمیه، سن، رتبه کنکور، ماه‌های سپری شده از شروع تحصیل در دانشجویان دکترای حرفه‌ای در بین دو گروه MCQ و EMQ تفاوت آماری معنی دار نداشت (جدول ۱). .

جدول ۱: خصوصیات زمینه‌ای شرکت‌کنندگان در مطالعه

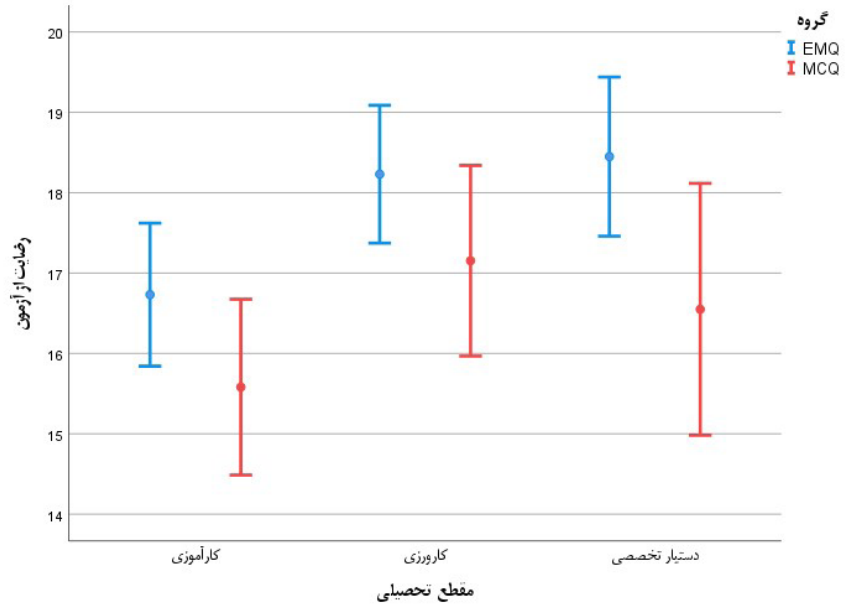
ویژگی	زیرویژگی	نوع آزمون*		P-Value
		MCQ	EMQ	
مشارکت کننده	دستیار بالینی کارورز کارآموز	۲۰ (٪ ۲۲/۵) ۲۶ (٪ ۲۹/۲) ۴۳ (٪ ۴۸/۳)	۲۰ (٪ ۲۳) ۲۶ (٪ ۲۹/۹) ۴۱ (٪ ۴۷/۱)	۰/۹۸۸
جنسیت	زن مرد	۵۹ (٪ ۶۶/۳) ۳۰ (٪ ۳۳/۷)	۵۷ (٪ ۶۵/۵) ۳۰ (٪ ۳۴/۵)	۰>۰/۹۹
سن	-	۲۸/۵ ± ۴/۵	۱۲۵/۱ ± ۴/۴	۰/۵۷۸
سهمیه قبولی	آزاد سهمیه	۵۷ (٪ ۶۴/۰) ۳۲ (٪ ۳۶/۰)	۵۱ (٪ ۵۸/۶) ۳۶ (٪ ۴۱/۴)	۰/۵۳۶
رتبه کنکور	-	۵۱۰۶/۶ ± ۶۴۶/۹	۹۴۷۱ ± ۳۵۴/۸	۰/۴۲۶
ماههای سپرس شده از تحصیل	-	۵۱/۷ ± ۱۹/۶	۵۱/۶ ± ۲۰/۵	۰/۹۷۴
گروه آموزشی	زنان و زایمان بیماری‌های داخلی	۲۹ (٪ ۳۲/۶) ۶۰ (٪ ۶۷/۴)	۲۷ (٪ ۳۱/۰) ۶۰ (٪ ۶۹/۰)	۰/۸۷۲

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

نمره آزمون EMQ (۲/۵±۱۷/۱) بیشتر از آزمون MCQ (۲/۶±۱۶/۶) بود اما تفاوتی از نظر آماری وجود نداشت (۰,۱۸=□) با این حال رضایت دانشجویان از آزمون EMQ بیشتر از آزمون MCQ بود (۲/۵±۱۷/۵ در مقابل ۳/۳±۱۶/۲ و P=۰/۰۰۴). همانطور که نمودار ۱ نشان می دهد هر چند در مقطع کارآموزی و دستیاری نمره آزمون EMQ بالاتر از آزمون MCQ بود اما این تفاوت از نظر آماری معنی دار نبود. در خصوص رضایت از آزمون (نمودار ۲) نیز در هر سه مقطع تحصیلی رضایت از آزمون EMQ بیشتر از آزمون MCQ بود اما تنها در مقطع دستیاری این تفاوت از نظر آماری معنی دار بود (P=۰/۰۴۰). در نهایت نمره آزمون EMQ با رتبه کنکور همبستگی معکوس و معنی داری داشت (r= -۰/۳۹۹، P=۰/۰۱۱) اما این همبستگی برای آزمون MCQ مشاهده نشد (r= -۰/۲۵۸، P=۰/۰۹۵).



نمودار ۱: مقایسه نمره آزمون در دو گروه مورد مطالعه به تفکیک مقطع تحصیلی(میله های خطا=دامنه اطمینان ۹۵ درصدی)



نمودار ۲: مقایسه رضایت از آزمون در دو گروه مورد مطالعه به تفکیک مقطع تحصیلی (میله های خطا=دامنه اطمینان ۹۵ درصدی)

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

بحث و نتیجه‌گیری:

روش‌های ارزشیابی پایانی در حوزه پزشکی، غالباً بر آزمون‌های چندگزینه‌ای متکی هستند که عمدتاً به سنجش یادآوری اطلاعات محدود می‌شوند. این در حالی است که نیاز به ارزیابی جامع‌تر دانشجویان پزشکی، به ویژه در سنجش توانایی آن‌ها در کاربرد دانش در موقعیت‌های واقعی، امری ضروری است [۹،۱۰]. نظیم و همکاران در سال ۲۰۱۹ در پژوهشی با عنوان "ارزیابی مهارت‌های استدلال بالینی با استفاده از آزمون اسکریپت تطابق (SCT) و سؤالات تطبیق گسترده: آزمایشی برای کارآموزان اورولوژی" دریافتند که اکثر شرکت‌کنندگان از آزمون جورکردنی گسترده پاسخ احساس رضایت داشتند و همچنین می‌توان برای ارزیابی و تقویت مهارت‌های استدلال بالینی مورد استفاده قرار گیرند [۱۹]. همسو با مطالعه مذکور، نتایج مطالعه حاضر نیز نشان می‌دهد که آزمون EMQ توانسته است رضایت دانشجویان را کسب کند و همچنین از نظر ارزیابی دانشجو نسبت به آزمون MCQ بهتر عمل کرده است.

فری و همکاران در سال ۲۰۲۲ در پژوهشی با عنوان "آمار قالب آیتم و خوانایی سوالات تطبیق گسترده به عنوان ابزاری موثر برای ارزیابی دانشجویان پزشکی" دریافتند که آزمون EMQ برای ارزیابی دانشجویان پزشکی کارآمدتر و مناسب‌تر می‌باشد و ویژگی‌های قدرتمند قابلیت اطمینان را نشان می‌دهد [۴]. همچنین در مطالعه حاضر نیز نتایج به دست آمده نمره ارزیابی توسط آزمون EMQ را بالا نشان می‌دهد که بیانگر ارزیابی بهتر دانشجویان می‌باشد. از طرفی بالا بودن میزان رضایت دانشجویان از این آزمون می‌تواند بیان کننده‌ی این امر باشد که آزمون توانسته است مهارت‌های مرتبه بالاتر را بهتر ارزیابی کند و توانایی تشخیص دانشجویان توانمندتر را نیز داشته باشد.

در مطالعه‌ی باکتا و همکاران در سال ۲۰۰۵ با عنوان "استفاده از تئوری پاسخ آیتم برای کشف ویژگی‌های روان‌سنجی آزمون سوالات تطبیق گسترده در آموزش پزشکی در مقطع کارشناسی" نتایج نشان می‌دهد که اگر سوالات و پاسخ‌های آزمون جورکردنی گسترده پاسخ به درستی طراحی شوند می‌تواند دانشجویان توانمندتر را بهتر تشخیص دهد و از طرفی بیان می‌کند که جنسیت نیز می‌تواند بر پاسخ دانشجویان به عوامل حواس‌پرتی تأثیر بگذارد [۶]. از طرفی در مطالعه ما نتایج نشان داد که جنسیت در پاسخ‌دهی به سوالات تاثیر معناداری ندارد. ایسوگلز و همکاران در سال ۲۰۱۳ در پژوهشی با عنوان "سؤالات چندگزینه‌ای برای شناسایی دانشجویان پزشکی و علوم زیست پزشکی که عملکرد ضعیفی دارند، نسبت به سؤالات منطبق گسترده برتری دارند." دریافتند که آزمون MCQ نسبت به آزمون EMQ عملکرد بهتری در شناسایی دانشجویان ضعیف دارد [۷]. این در حالیست که در مطالعه باکتا و همکاران بیان می‌کند آزمون EMQ نسبت به آزمون MCQ در ارزیابی دانشجویان توانمندتر بهتر عمل می‌کند [۶]. همسو با مطالعه مذکور، در مطالعه حاضر نیز نتایج نشان می‌دهد نمره آزمون EMQ بیشتر از آزمون MCQ بوده که به معنی ارزیابی بهتر دانشجویان می‌باشد و همچنین نمره آزمون EMQ با رتبه کنکور همبستگی معکوس و معنی‌داری داشت و به این معنی می‌باشد هرچه رتبه کنکور بالاتر باشد نمره آزمون EMQ دانشجویان پایین‌تر است که می‌توان نتیجه گرفت دانشجویان با عملکرد پایین در کنکور، نمره آزمون EMQ پایین نیز کسب می‌کنند. و از طرفی دانشجویان نسبت به آزمون EMQ رضایت بیشتری داشتند.

مطالعه نظیم و همکاران در سال ۲۰۱۹ نتایج بیان می‌کند که آزمون EMQ برای ارزیابی مهارت‌های مرتبه بالاتر مانند استدلال بالینی مناسب‌تر است [۱۹]. این در حالیست که مطالعه فام و همکاران در سال ۲۰۱۸ با عنوان "انتخاب ارزیابی های پزشکی: آیا سوال چندگزینه ای باعث نمره می شود؟" دریافتند که آزمون MCQ می‌تواند مهارت‌های مرتبه بالاتر رابه خوبی ارزیابی کند [۲۰]. در مطالعه حاضر نیز نتایج میزان ارزیابی بهتر آزمون EMQ را در دانشجویان نشان می‌دهد. مک‌کوبری در سال ۲۰۰۴ در پژوهشی تحت عنوان "بهبود عادلانه بودن سوالات چند گزینه‌ای: مرور ادبیات" بیان می‌کند که فراگیر بودن آزمون‌های MCQ از کارایی و در نتیجه پایایی آنها ناشی می‌شود و می‌تواند سطوح پایین دانش را به خوبی ارزیابی کند، اما با این حال آزمون

EMQ برای ارزیابی‌های مهارت‌های عملی منصفانه‌ترین قالب هستند [۲۱]. همچنین نتایج پژوهش حاضر نیز چنین نشان می‌دهد که دانشجویان از آزمون EMQ رضایت بیشتری داشته‌اند که می‌تواند به دلیل این امر باشد که این آزمون بهتر توانسته است مهارت‌های بالینی دانشجویان را ارزیابی کند که میتوان از نمره بهتر آزمون EMQ نسبت به آزمون MCQ نیز این مورد را دریافت. لذا توصیه می شود در مطالعاتی که در حوزه طراحی آزمون های جدید در رشته های مختلف می باشد، شواهد حاصل از مطالعه حاضر نیز مدنظر پژوهشگران باشد [۲۲].

نتیجه‌گیری:

در مجموع علیرغم اینکه در حوزه مقایسه روش های تدریس مطالعات متعددی صورت گرفته است [۳۰-۲۳]. اما مطالعات در حوزه روش های ارزیابی محدودتر و عمدتا مروری بوده و بر پایه مقایسه دو روش ارزیابی نیز نمی باشند [۳۳-۳۱]. مطالعات حاضر با هدف پوشش این خلا دانش صورت گرفت و با توجه به نتایج به دست آمده در این پژوهش می توان چنین توصیه کرد که آزمون جورکردنی گسترده پاسخ (EMQ) به دلیل توانایی بالا در میزان ارزیابی مهارت‌های مرتبه بالاتر و همچنین میزان رضایتمندی دانشجویان از این نوع آزمون در سنجش دانشجویان در محیط‌های بالینی استفاده شود.

از محدودیت‌های این مطالعه می توان به این مساله اشاره کرد که دستیاران تخصصی تنها از دو گروه آموزشی وارد مطالعه شده بودند. هر چند این دو گروه از گروه های ماژور هستند، با این حال افزایش تنوع گروه های آموزشی می تواند منجر به افزایش قدرت تعمیم پذیری نتایج شود. از سوی دیگر، استفاده از سه مقطع مختلف تحصیلی (کارآموزی، کارورزی و دستیار تخصصی) از نقاط قوت مطالعه می باشد که نمای نسبتا کاملی از میزان تاثیر دو روش در مقاطع مختلف تحصیلی فراهم می کند. برای تکمیل تصویر مقایسه این دو روش ارزیابی دانشجویی، طراحی مطالعات در سایر رشته های علوم پزشکی توصیه می شود.

ملاحظات اخلاقی:

پیروی از اصول اخلاق در پژوهش

این مطالعه در کمیته اخلاق سازمانی IR.MUMS.REC.1400.232 دانشگاه علوم پزشکی مشهد تصویب شده است.

حامی مالی

این مقاله با حمایت مالی دانشگاه علوم پزشکی مشهد انجام شده است.

سهم نویسندگان

کلیه نویسندگان به طور مساوی در تدین مقاله مشارکت داشته‌اند.

تعارض منافع

هیچ گونه تعارض منافعی در پژوهش حاضر وجود ندارد.

تشکر و قدردانی

از حمایت‌های معاونت پژوهشی دانشگاه علوم پزشکی مشهد در اجرای طرح حاضر (کد طرح ۹۸۰۷۵۶) قدردانی می شود. همچنین از همکاری مرکز تحقیقات آموزش علوم پزشکی دانشگاه علوم پزشکی مشهد قدردانی می‌شود.

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

References

1. Vanderbilt AA, Feldman M, Wood IK. Assessment in undergraduate medical education: A review of course exams. Med Educ Online. 2013; 18:1-5. DOI: 10.3402/meo.v18i0.20438
2. Grainger R, Dai W, Osborne E, Kenwright D. Medical students create multiple-choice questions for learning in pathology education: A pilot study. BMC Med Educ. 2018; 18(1):201. DOI: 10.1186/s12909-018-1312
3. Vyas R, Supe A. Multiple choice questions: A literature review on the optimal number of options. Natl Med J India. 2008; 21(3):130-3. PMID: 19004145
4. Frey A, Leutritz T, Backhaus J, Hörnlein A, König S. Item format statistics and readability of extended matching questions as an effective tool to assess medical students. Sci Rep. 2022; 12(1):20982. DOI: 10.1038/s41598-022-25481-y
5. Aljarallah BM. Evaluation of modified essay questions (MEQ) and multiple choice questions (MCQ) as a tool for assessing the cognitive skills of undergraduate medical students. Int J Health Sci. 2011; 5(1):39. PMID: 22489228
6. Bhakta B, Tennant A, Horton M, Lawton G, Andrich D. Using item response theory to explore the psychometric properties of extended matching questions examination in undergraduate medical education. BMC Med Educ. 2005; 5(1):9. DOI: 10.1186/1472-6920-5-9
7. Eijsvogels TMH, van den Brand TL, Hopman MTE. Multiple choice questions are superior to extended matching questions to identify medicine and biomedical sciences students who perform poorly. Perspect Med Educ. 2013; 2(5-6):252-263. DOI: 10.1007/s40037-013-0068-x
8. Karshki H, Abadi H, Bagheri M, Ghanbari Moghadam A. A review of the basics of determining the cutting point in academic achievement tests and the introduction of Angoff scientific method. Horiz Med Educ Dev. 2022; 12(4):96-85. [In Persian] DOI: 10.22038/hmed.2021.54426.1116
9. Scouller K. The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. High Educ. 1998; 35(4):453-72. DOI: 10.1023/A:1003196224280
10. Banning M. A review of clinical decision making: Models and current research. J Clin Nurs. 2008; 17(2):187-95. DOI: 10.1111/j.1365-2702.2006.01791.x
11. 11- Fenderson BA, Damjanov I, Robeson MR, Veloski JJ, Rubin E. The virtues of extended matching and uncued tests as alternatives to multiple choice questions. Hum Pathol. 1997; 28(5):526-32. DOI: 10.1016/s0046-8177(97)90073-3
12. Kreiter CD, Ferguson K, Gruppen LD. Evaluating the usefulness of computerized adaptive testing for medical in-course assessment. Acad Med. 1999; 74(10):1125-8. DOI: 10.1097/00001888-199910000-00016
13. Case SM, Swanson DB. Extended-matching items: A practical alternative to free-response questions. TLM. 1993; 5(2):107-15. DOI: 10.1080/10401339309539601
14. Chandratilake M, Davis M, Ponnamperuma G. Assessment of medical knowledge: the pros and cons of using true/false multiple choice questions. Natl Med J India. 2011; 24(4):225-8. PMID: 22208143
15. Beullens J, Struyf E, Van Damme B. Do extended matching multiple-choice questions measure clinical reasoning? Med Educ. 2005; 39(4):410-7. DOI: 10.1111/j.1365-2929.2005.02089.x
16. Lukić IK, Gluncić V, Katavić V, Petanjek Z, Jalsovec D, Marusić A. Weekly quizzes in extended-matching format as a means of monitoring students' progress in gross anatomy. Ann Anat. 2001; 183(6):575-9. DOI: 10.1016/S0940-9602(01)80073-X
17. Case SM, Swanson DB. Evaluating diagnostic pattern: A psychometric comparison of items with 15, 5 and 2 options. Inmeeting of the American Educational Research Association. San Francisco; 1989.
18. Duthie S, Hodges P, Ramsay I, Reid W. EMQs: A new component of the MRCOG Part 2 exam. TOG. 2006; 8(3):181-5. DOI: 10.1576/toag.8.3.181.27255
19. Nazim SM, Talati JJ, Pinjani S, Biyabani SR, Ather MH, Norcini JJ. Assessing clinical reasoning skills using Script Concordance Test (SCT) and extended matching questions (EMQs): A pilot for urology trainees. J Adv Med Educ Prof. 2019; 7(1):7-13. DOI: 10.30476/

Downloaded from dsme.hums.ac.ir at 13:22 +0330 on Thursday November 5th 2026

JAMP.2019.41038

20. Pham H, Trigg M, Wu S, O'Connell A, Harry C, Barnard J, et al. Choosing medical assessments: Does the multiple-choice question make the grade? Educ Health (Abingdon). 2018; 31(2):65-71. DOI: 10.4103/efh.EfH_229_17

21. McCoubrie P. Improving the fairness of multiple-choice questions: A literature review. Med Teach. 2004; 26(8):709-12. DOI: 10.1080/01421590400013495

22. Asadinezhad M, Khademi S, Ziaei AH, Olumi S. Implementation of pre-internship exam in the radiology technology field of Mashhad University of Medical Sciences. Horiz Med Educ Dev. 2023; 14(4):64-74.[In Persian] DOI: 10.22038/HMED.2023.70704.1259

23. Milani N, Khadem-Rezaiyan M, Ataee Z. Comparison of learning in flipped classroom and traditional classroom in family medicine assistants. Horiz Med Educ Dev. 2024; 15(3):6-17. [In Persian] DOI: 10.22038/HMED.2023.67415.1231

24. Maleki S, Ahmadi S, Habibi A, Seidi M, Tehrani Far F. Comparing the effect of two teaching approaches in lecture and multimedia on nurses' problem solving skills. Horiz Med Educ Dev. 2024. [In Persian] DOI: 10.22038/hmed.2024.78813.1354

25. Abbaszadeh Mehrabadi E, Gheshlaghi M, Barzegar S, Javadi M, Nasiriani K, Nazmiah H, et al. Comparing the effectiveness of basic cardiopulmonary resuscitation training using. Horiz Med Educ Dev. 2024; 15(3):84-94. [In Persian] DOI: 10.22038/HMED.2024.77980.1331

26. Ebrahimi Shahroodi O, Manzari ZS, Aghebati N, Esmaily H. Comparison of the effect of two educational Methods of. Horiz Med Educ Dev. 2024; 15(2):1-10. [In Persian] DOI: 10.22038/HMED.2024.70413.1256

27. Sargolzaei F, Ghadami A, Omid A, Azarmehr T. The Effect of training using the combination of virtual reality and augmented reality in coronary artery surgery on the learning rate of undergraduate students in the operating room. Horiz Med Educ Dev. 2023; 14(4):43-52. [In Persian] DOI: 10.22038/HMED.2023.69629.1246

28. Najjari M, Moammeri F, Jamali J, Rezayan S. Evaluating the effect of two methods of virtual and traditional education on parasitology course scores of laboratory science students in Mashhad. Horiz Med Educ Dev. 2023; 14(2):12. [In Persian] DOI: 10.22038/HMED.2022.65419.1221

29. Gerivani A, Behjati N, Montazer Ataei M, Karimi Monaghi H. Comparison of physiopathology stage training program by course method and national medical program, Case Study (Mashhad University of Medical Sciences). Horiz Med Educ Dev. 2022; 13(4):39-29. [In Persian] DOI: 10.22038/HMED.2022.60017.1173

30. Ghiyasvandian S, Haji Mohammad Hoseini M, Asayesh H, Sadeghi R, Azadeh M, Parvaresh Masoud M. The effect of team-based learning on students' learning and motivation of students of medical emergency in the triage lesson. Horiz Med Educ Dev. 2021; 12(1):34-43. [In Persian] DOI: 10.22038/HMED.2020.50081.1053

31. Rajabzadeh F, Milani N, Goshayeshi L. Education and exams of medical students in COVID pandemic. Horiz Med Educ Dev. 2020; 11(3):6-1. [In Persian] DOI: 10.22038/HMED.2020.50749.1066

32. Rajabpour M, Moonaghi HK. Effective tools and methods for evaluating the clinical performance of medical sciences students: A systematic review. Horiz Med Educ Dev. 2024; 15(2):69-81. [In Persian] DOI: 10.22038/HMED.2024.73743.1283

33. Keshmiri F, Heydari AS. Electronic evaluation methods: A scoping review. Horiz Med Educ Dev. 2022; 13(4):98-85. [In Persian] DOI: 10.22038/HMED.2021.58687.1176