

## Analytic Interpretation of Multiple Choice Questions in Embryology among Students in The College of Medicine, University of Mosul

Omar Riadh Al-Jalily\*

\*Department of Anatomy, College of Medicine, University of Mosul. Mosul, Iraq  
Correspondence: omarrh@uomosul.edu.iq

(Ann Coll Med Mosul 2025; 47 (1):73-78).

Received: 29<sup>th</sup> Sep. 2024; Accepted: 15<sup>th</sup> Oct. 2024.

### ABSTRACT

**Background:** Many methods are used to perform tests for students, but the best are those that achieve the highest level of deep and high learning for students and ease of analysis. One of the most important testing methods is multiple-choice questions (MCQs).

**Aim:** To analyze the multiple-choice questions in embryology subject for the students of the second class in the College of Medicine, University of Mosul, Iraq.

**Methods:** The multiple-choice method was used to conduct the final exam (first attempt/2022) for the embryology subject for second-year students at the College of Medicine in Mosul (697 students). The results were analyzed prospectively through a non-interventional observational study. The test included 100 multiple-choice tests, so each question had five choices (four wrong choices were used as distractions because they were close).

**Results:** Two tools are used to analyze each multiple-choice question: Difficulty Index and Discrimination Index. About 52 (52.0%) out of 100 multiple-choice questions were average (between 41-60), while about 46 (46.0%) were moderately easy (more than 61-80). About 2 (2.0%) were moderately difficult (between 21-40). In contrast, about 10 (10%) items were within acceptable values (0.2-0.29), while about 0 (0%) items were considered good (0.3-0.4), and about 0 (0%) were excellent (more than 0.4), about 1(%) was within the false zone (less than 0). The weak category (0-0.19) was seen in 89 (8.9%) students.

**Conclusion:** The test must use the Medium Difficulty Index and Good Discrimination Index items. Such an analysis could improve the quality of examinations in medical schools and review weak questions through the Discrimination index.

**Keywords:** MCQ, Difficulty index, and Discrimination index.

### التفسير التحليلي لأسئلة متعددة الخيارات في علم الأجنة بين طلبة كلية الطب جامعة الموصل

عمر رياض الجليلي\*

\*فرع التشريح ، كلية الطب ، جامعة الموصل ، الموصل ، العراق

### الخلاصة

**الخلفية:** هناك العديد من الطرق المستخدمة لإجراء الاختبارات للطلاب، ولكن أفضل هذه الطرق هي تلك التي تحقق أعلى مستوى من التعلم العميق والعالي للطلاب، بالإضافة إلى سهولة التحليل. ومن أهم طرق الاختبار أسئلة متعددة الخيارات MCQs.

**الأهداف:** تحليل أسئلة متعددة الخيارات من متعدد في مادة علم الأجنة لطلاب السنة الثانية في كلية الطب جامعة الموصل، العراق.

**الطرق:** تم استخدام طريقة الأسئلة متعددة الخيارات لإجراء الامتحان النهائي (المحاولة الأولى / ٢٠٢٢) لمادة علم الأجنة لطلاب السنة الثانية في كلية الطب بالموصل (٦٩٧ طالباً). تم تحليل النتائج بشكل استباقي من خلال دراسة مراقبة غير تدخلية. تضمن الاختبار ١٠٠ اختبار متعدد الخيارات من متعدد، بحيث يكون هناك خمسة خيارات لكل سؤال (تم استخدام أربعة خيارات خاطئة لأنها متقاربة).

**النتائج:** تم استخدام أداتين لتحليل كل سؤال من أسئلة متعددة الخيارات: مؤشر الصعوبة ومؤشر التمييز. وكان حوالي ٥٢ (٥٢.٠%) من أصل ١٠٠ سؤال من أسئلة متوسطة (بين ٤١-٦٠)، في حين كان حوالي ٤٦ (٤٦.٠%) سهلة إلى حد ما (أكثر من ٦١-٨٠). حوالي ٢ (٢.٠%) كانت صعبة إلى حد ما (بين ٢١-٤٠). وعلى النقيض من ذلك، كان حوالي ١٠ (١٠.٠%) من العناصر ضمن القيم المقبولة (٠.٢-٠.٢٩)، بينما تم اعتبار حوالي ٠ (٠.٠%) من العناصر جيدة (٠.٣-٠.٤)، وحوالي ٠ (٠.٠%) كانت ممتازة (أكثر من ٠.٤)، وحوالي ١ (١.٠%) كانت ضمن المنطقة الخاطئة (أقل من ٠). لوحظت الفئة الضعيفة (٠-٠.١٩) في ٨٩ (٨.٩%) طالبًا.

**الاستنتاجات:** إن استخدام مؤشر الصعوبة والمتوسطة ومؤشر التمييز الجيد ضروري في الاختبار. يمكن لمثل هذا التحليل تحسين جودة الامتحانات في كليات الطب ومراجعة الأسئلة الضعيفة من خلال مؤشر التمييز.

**الكلمات المفتاحية:** أسئلة متعددة الخيارات; مؤشر الصعوبة; مؤشر التمييز.

## INTRODUCTION

Since medical students will eventually become doctors, assessing their knowledge about various medical issues is necessary to evaluate their learning abilities in these areas <sup>1,2</sup>. These students may learn deeply through multiple approaches, such as oral and written. In addition to the objective structured practical/clinical examination, known as OSPE and OSCE, there are two formats for the written type: questions with multiple choices and essay questions (long, short, and modified). However, the evaluation may also help the learner develop an appropriate perspective on many medical topics <sup>3,4</sup>. Since these students are among the most significant stakeholders, it is necessary to conduct several research that concentrates on the methods of their exams. Each year, thousands of students join the medical colleges at various Iraqi universities <sup>4</sup>. Because they are easily evaluated and can assess students' deep learning and cognition, multiple choice questions, or MCQs, are the most popular questions <sup>5,6</sup>. It is also better because it is authentic, reliable, and simple to give a score <sup>7,8</sup>. The examination procedures, particularly multiple-choice questions, require study and interpretation to provide an appropriate strategy for evaluating medical students' knowledge. This is also the most significant way to prevent these students from being given inappropriate or invalid questions. This study aims to examine the multiple-choice questions in embryology for the students in the second class at the University of Mosul, Iraq's College of Medicine, as there haven't been many analyses like this in our area.

## Experimental Method

An analysis of the multiple-choice questions on embryology for a final examination/first attempt of 697 students in 2022 was conducted prospectively through a Biometric descriptive design at the College of Medicine, University of Mosul.

One hundred multiple-choice questions with five possible answers were provided (four distractions representing erroneous answers and one correct item). Each multiple-choice question is examined using the level of difficulty Index and Discrimination Index (Tables 1 and 2). Chi-square analysis was used for the statistical analysis.

The exam took three hours to complete, with each question worth one mark and no deductions for incorrect answers. It was drafted on paper, and each student used an answer sheet prepared by seasoned embryology lecturers. Two sets of answers were used to prevent copying between nearby medical students. The answer sheets were gathered to use Microsoft Excel 2020 for evaluation. Three groups composed the whole group: the middle group 1/3, the lower 1/3 (lower ability group-LAG), and the higher 1/3 (higher ability group -HAG).

The analysis employed data from the higher and lower-ability groups based on the necessity of computing the indexes <sup>9-11</sup>.

The following equations are applied:

1. Index of Difficulty =  $[(H + L)/N] \times 100$ . Difficulty Index values range from 0% to 100%, with 70% too easy <sup>12</sup>.
2.  $2 \times [(H - L)/N] =$  Discrimination Index. Excellent Discrimination Index Value 0.25, on a scale of 0 to 1.

**N** is the total number of pupils in the groups classified as upper and lower third.

**H** is the number of students in the higher ability group who correctly answered the item.

**L** is the proportion of students in the lower ability group who correctly answered the item <sup>13</sup>

Table 1 The range of the study's difficulty index.

Cutoff value	Interpretation
0 – 20	Very difficult
21 – 40	Moderately difficult
41-60	Average
61-80	Moderately easy
81-100	Very easy

Table 2 The range of the study's discrimination index.

Cutoff value	Interpretation
Less than 0	Faulty question
0 - 0.19	Poor discrimination
0.2 - 0.29	Acceptable discrimination
0.3 - 0.4	Good discrimination
More than 0.4	Excellent question

## RESULTS

The multiple-choice question designs in this study are presented understandably and logically. The exam was administered to around 697 students, with an 88.5 percent success rate, of which about 617 passed. Actually, medical students' performance can be enhanced by encouraging study.

The project aimed to get the correct answer for multiple-choice questions by using item analysis. Of the 100 multiple-choice questions, 52 (52.0%) were average (between 41-60), 46 (46.0%) were moderately easy (between 61-80), and 2 (2.0%) were moderately difficult (between 21-40) (Table 3).

Conversely, almost 10% of the items fell within the acceptable range of 0.2 to 0.29, 0% of the items were rated as good (0.3 - 0.4), 0% of the items were rated as exceptional (more than 0.4), and 1% of the items were in the false region (less than 0). Eighty-nine students (8.9%) fell into the poor category (0 - 0.19) (Table 4).

Table 3 Index for difficulty for the study data

Cutoff value	Interpretation	Number of items
0 -20	Very difficult	--
21-40	Moderately difficult	2
41-60	Average	52
61-80	Moderately easy	46
81-100	Very easy	--

Table 4. The index for discrimination of the study's data.

Cutoff value	Interpretation	Number of items
Less than 0	Faulty question	1
0 - 0.19	Poor discrimination	89
0.2 - 0.29	Acceptable discrimination	10
0.3 - 0.4	Good discrimination	0
More than 0.4	Excellent question	0

## DISCUSSION

The assessment and analysis of the multiple-choice questions were crucial for several reasons, including how it affects medical students' learning quality and makes it possible to conduct a relevant and precise examination that clarifies the knowledge and abilities critical to the medical industry<sup>14</sup>.

The success of the students' exam performance served as the cornerstone of their education. Since assessments rely on factual information, students must be provided with a straightforward path or strategy<sup>15</sup>. On the other hand, a thorough approach was required for exams requiring a high degree of cognitive ability. Thus, the leader of these students and their learning method is the manner of assessment<sup>16-18</sup>.

Much research has been done on this subject, particularly in our community. Therefore, professors are motivated to raise the standard of these tests.

The data analysis for this work showed that, on average, 52 out of 100 multiple-choice questions

(MCQs) were simple, 46 were moderately easy, and two were challenging. Conversely, regarding the Discrimination Index, almost 10% of the products fell within the acceptable range, no good or exceptional items were discovered, and approximately 89% fell into the poor group. One item was thought to be defective. Not much research has been published on medical students' embryology knowledge. Rao et al.<sup>19</sup> used fifty multiple-choice questions to conduct an item analysis on 100 medical students studying anatomy. They discovered that 31 (62%) of the items had difficulty indices in the acceptable range but that 16 (32%) and 3 (6%) of the items were too easy and challenging, respectively<sup>18</sup>.

The results of this study are consistent with those of Patel, who researched medical students studying microbiology. Patel reported that 12 (30%) of the items on the difficulty index were in the ideal range, 18 (45%) were in an acceptable range, but only 7 (17.5%) of the items were easy, and 3 (7.5%) were difficult<sup>20</sup>. Further research on biochemistry students studying medicine revealed that, of the 30 multiple-choice questions, 21 had an adequate difficulty level, one was too easy, and eight were excessively tough. 8 (26.67%) products had a discrimination index in the recommended range, sixteen (53.33%) in the acceptable range, and 6 (20%) were classified as poor. The variation in sample size may cause a discrepancy with our work<sup>20,21</sup>. Additionally, another study (Kolte)<sup>22</sup> found that, in line with this work, the P values of 26 (65%) items were within an acceptable range, 10 (25%) things were easy, and 4 (10%) items were challenging. Nonetheless, the 60% item discrimination index was excellent. Also, in line with this work, Gajjar et al.<sup>23</sup> found that out of 50 items, 24 had a "good to excellent" difficulty index, whereas 15 also had a good to excellent discrimination index.

Revisions to the medical curriculum are necessary to achieve the primary objective of any medical college, which is to produce medical practitioners with reflective and self-directed personalities through the three modes of learning (deep, surface, and strategic)<sup>24</sup>. In actuality, surface learning relies on memorization rather than in-depth analysis and recognition of the material, and it is influenced by failure-related fretting. Conversely, the in-depth ones depend on the fascinating intricacies of the subjects and the logical relationship between the current and outdated knowledge about each topic.

If properly constructed, multiple choice questions allow lecturers to assess many students' deep learning and solid cognitive abilities over a broad range of knowledge and discriminate between high and low quality<sup>25-27</sup>. However, more time and effort

are required to complete these multiple-choice problems. Thus, it is essential to analyze multiple-choice questions to determine their suitability. Using a difficulty index to identify objects that are too easy, too challenging, and recommended is one way to arrive at this conclusion. About 52% of the multiple-choice questions in this work have been recommended items based on the difficulty index<sup>28</sup>.

Overly simple questions are meant to warm up students' abilities. In contrast, challenging questions are intended to assess students' high talents but don't help differentiate between students who achieve well and those who don't<sup>29</sup>.

It is crucial to adjust the wording and content of every question to fall within the recommended range to ensure that it is appropriate for both high and low-ability groups. According to the present research, around 46% of the items require modification to be placed into the recommended class using the difficulty index (in language and text). However, 10% of the study's multiple-choice questions fell within the acceptable quality range determined by the discrimination index. The discrimination index can distinguish between different student quality levels. The discrimination and difficulty indexes are related to each other. Mehta et al.<sup>8</sup> did an excellent job with difficulty and discrimination index in this work. 74% "recommended" multiple-choice questions (MCQs out of 100%) according to the difficulty index and 34% "discrimination index," and this is in line with our study.

There are five choices for each question pertaining to our work: A, B, C, D, and E. Analysis of works spanning more than 80 years, however, focused more on the quality of decisions than the quantity of them<sup>30,31</sup>.

This study's strength lies in the appropriate number of questions (n = 100) to achieve valuable results.

## CONCLUSIONS

Using the average difficulty index and good discriminating index items in the upcoming exam is necessary. By using a discriminating index, this type of analysis can help medical colleges enhance the quality of their exams and correct poorly written questions.

It is crucial to create a question bank because of the significant correlation that was discovered between the two indices (difficulty and discrimination). Additional indexes, such as distractor indices, should be included to improve the quality of the questions.

## Funding

This work was done with self-supporting.

## Conflict of Interest:

There is no conflict of interest in this work.

## REFERENCES

1. Saeed GT, Hassan AE, Al Omary HL, Alawad ZM. Multiple choice and essay questions assess success rates in medical physiology. *Journal of the Faculty of Medicine Baghdad*. 2017;59(4):333-5.
2. Rust C, O'donovan B, Price M. A social constructivist assessment process model: how the research literature shows us this could be best practice. *Assessment & Evaluation in Higher Education*. 2005 Jun 1;30(3):231-40. DOI: 10.1080/02602930500063819
3. Sood R, Singh T. Assessment in medical education: Evolving perspectives and contemporary trends. *Natl Med J India*. 2012 Nov 1;25(6):357-64.
4. Vegi VA, Sudhakar PV, Bhimarasetty DM, Pamarthi K, Edara L, Kutikuppala LS, Suvvari TK, Anand S. Multiple-choice questions in assessment: Perceptions of medical students from a low-resource setting. *Journal of Education and Health Promotion*. 2022 Jan 1;11(1):103. DOI: 10.4103/jehp.jehp\_621\_21
5. Abdel-Hameed AA, Al-Faris EA, Alorainy IA, Al-Rukban MO. The criteria and analysis of good multiple choice questions in a health professional setting. *Saudi Medical Journal*. 2005 Oct 1;26(10):1505-10.
6. Case SM, Swanson DB. Constructing written test questions for the basic and clinical sciences. Philadelphia: National Board of Medical Examiners; 1998 Oct 20.
7. Tarrant M, Ware J. A framework for improving the quality of multiple-choice assessments. *Nurse Educator*. 2012 May 1;37(3):98-104. DOI: 10.1097/NNE.0b013e31825041d0
8. Mehta M, Banode S, Adwal S. Analysis of multiple choice questions (MCQ): important part of assessment of medical students. *International Journal of Medical Research and Review*. 2016;4:199-204.
9. Hingorjo MR, Jaleel F. Analysis of one-best MCQs: the difficulty index, discrimination index and distractor efficiency. *JPMA-Journal of the Pakistan Medical Association*. 2012 Feb 1;62(2):142.
10. Singh T, Gupta P, Singh D. Principles of medical education. Jaypee Brothers Medical Publishers; 2020 Oct 7.
11. Cizek GJ, O'Day DM. Further investigation of nonfunctioning options in multiple-choice test items. *Educational and Psychological Measurement*. 1994 Dec;54(4):861-72. DOI: 10.1177/0013164494054004002
12. P. Date, A. S. Borkar, R. T. Badwaik, R. A. Siddiqui, T. R. Shende, and A. V. Dashputra, "Item analysis as tool to validate multiple choice question bank in pharmacology," *Int J Basic Clin Pharmacol*, vol. 8, no. 9, 2019, doi: 10.18203/2319-2003.ijbcp20194106.
13. Garg R, Kumar V, Maria J. Analysis of multiple choice questions from a formative assessment of medical students of a medical college in Delhi, India. *Int J Res Med Sci*. 2019 Jan;7(1):174-7. DOI: 10.18203/2320-6012.ijrms20185375
14. Tarrant M, Ware J. A framework for improving the quality of multiple-choice assessments. *Nurse Educator*. 2012 May 1;37(3):98-104. DOI: 10.1097/NNE.0b013e31825041d0
15. Scouller K. The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher education*. 1998 Jun;35(4):453-72. DOI: 10.1023/A:1003196224280
16. Trigwell K, Prosser M. Improving the quality of student learning: the influence of learning context and student approaches to learning on learning outcomes. *Higher education*. 1991 Oct;22(3):251-66. DOI: 10.1007/BF00132290
17. Biggs J, Tang C. Teaching for quality learning at university. Buckingham. 2003.
18. Reid WA, Duvall E, Evans P. Relationship between assessment results and approaches to learning and studying in year two medical students. *Medical education*. 2007 Aug;41(8):754-62. DOI: 10.1111/j.1365-2923.2007.02801.x
19. Rao C, Kishan Prasad HL, Sajitha K, Permi H, Shetty J. Item analysis of multiple choice questions: Assessing an assessment tool in medical students. *Int J Educ Psychol Res*. 2016 Oct 1;2(4):201-4. DOI: 10.4103/2395-2296.189670
20. Patel RM. Use of Item analysis to improve quality of Multiple Choice Questions in II MBBS. *Journal of Education Technology in Health Sciences*. 2017 Apr 29;4(1):22-9. DOI: 10.18231/2393-8005.2017.0007

21. Saxena S, Srivastava PC, Mallick AK, Joshi HS, Das B. Item analysis: An unmatched tool for validating MCQs in Medical Education. *Indian J Basic Appl Med Res.* 2016 Sep;5(4):263-69.
22. Kolte V. Item analysis of multiple choice questions in physiology examination. *Indian Journal of Basic and Applied Medical Research.* 2015;4(4):320-6.
23. Gajjar S, Sharma R, Kumar P, Rana M. Item and test analysis to identify quality multiple choice questions (MCQs) from an assessment of medical students of Ahmedabad, Gujarat. *Indian Journal of Community Medicine.* 2014 Jan 1;39(1):17-20. DOI: 10.4103/0970-0218.126347
24. Evans C, Waring M. Applications of styles in educational instruction and assessment. *Handbook of intellectual styles: Preferences in cognition, learning, and thinking.* 2012:295-327.
25. Downing SM. Assessment of knowledge with written test forms. *International handbook of research in medical education.* 2002:647-72.
26. McCoubrie P. Improving the fairness of multiple-choice questions: a literature review. *Medical teacher.* 2004 Dec 1;26(8):709-12.
27. Schuwirth LW, Van Der Vleuten CP. Different written assessment methods: what can be said about their strengths and weaknesses?. *Medical education.* 2004 Sep;38(9):974-9. DOI: 10.1111/j.1365-2929.2004.01916.x
28. Farley JK. The multiple-choice test: writing the questions. *Nurse Educator.* 1989 Nov 1;14(6):10-2.
29. Haladyna TM, Downing SM. How many options is enough for a multiple-choice test item?. *Educational and psychological measurement.* 1993 Dec;53(4):999-1010. DOI: /10.1177/0013164493053004013
30. Rodriguez MC. Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational measurement: issues and practice.* 2005 Jun;24(2):3-13. DOI: 10.1111/j.1745-3992.2005.00006.x
31. Glaser R, Chudowsky N, Pellegrino JW, editors. *Knowing what students know: The science and design of educational assessment.* National Academies Press; 2001 Oct 27.