

An Analysis of the 1993 Comprehensive Medical Examination Multiple Choice Questions*

Boonnart Laisnitsarekul**

KitpramukTantayaporn*** Jiruth Sriratanaban****

Laisnitsarekul B, Tantayaporn K, Sriratanaban J. An Analysis of the 1993 Comprehensive Medical Examination Multiple Choice Questions. Chula Med J 1994 Aug; 38(8): 451-456

This descriptive research aimed to study the quality of the MCQ test, the 1993 Comprehensive Medical Examination, of the Faculty of Medicine, Chulalongkorn University in terms of item-difficulty index, item-discrimination index, reliability of test and determination of the number of excellent, good, revise, poor and negative questions by using the standard criteria of World Health Organization. The MCQ test was analysed by use of the CTIA item analysis program. The test consisted of 298 questions from 12 clinical departments: Obstetrics and Gynecology, Surgery, Pediatrics, Medicine, Psychiatry, Preventive and Social Medicine, Oto-laryngology, Radiology, Anesthesiology, Forensic Medicine, Ophthalmology and Orthopedic Surgery and Rehabilitation Medicine. The reliability of the test was 0.85. There were 21 excellent questions (7.05%), 32 good questions (10.74%), 20 revise questions (6.71%), 196 poor questions (65.77%) and 29 negative questions (9.73%). Both excellent and good questions were the most useful and should be kept in the question bank. Instructors who wrote the questions should completely check the questions, stem and distractors. The program director should have a policy of analysing the test every year for the sake of quality control.

Key words: *Comprehensive Medical examination, MCQ, Item analysis, Difficulty index, Discrimination index, Reliability.*

Reprint request: Laisnitsarekul B, Medical Education Unit, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand.

Received for publication. July 14, 1994.

* In the name of the 1993 Comprehensive Examination Committee.

** Medical Education Unit, Faculty of Medicine, Chulalongkorn University.

*** Department of Obstetrics and Gynecology, Faculty of medicine, Chulalongkorn University.

**** Department of Preventive and Social medicine, Faculty of Medicine, Chulalongkorn University.

บุญนาท ลายสนิทเสรีกุล, กิจประมุข ตันตยาภรณ์, จิรุตม์ ศรีรัตนบัลล์. การวิเคราะห์ข้อสอบปรนัย วิชาเวชศาสตร์ทั่วไป ประจำปีการศึกษา 2536. จุฬาลงกรณ์เวชสาร 2537 สิงหาคม; 38(8): 451-456

การวิจัยเชิงพรรณานี้ มีวัตถุประสงค์เพื่อศึกษาคุณภาพของข้อสอบปรนัย วิชาเวชศาสตร์ทั่วไป คณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่ใช้ในปีการศึกษา 2536 ในระดับความยากง่ายและอำนาจจำแนกของข้อสอบ ความเที่ยงของแบบสอบ ตลอดจนจำแนกจำนวนของข้อสอบที่ดีเยี่ยม ข้อสอบที่ดี ข้อสอบที่ควรได้รับการแก้ไข ข้อสอบที่ไม่ดีและข้อสอบเชิงลบ ตามเกณฑ์มาตรฐานขององค์การอนามัยโลก โดยนำข้อสอบดังกล่าวมาวิเคราะห์ด้วยโปรแกรม CTIA พบว่า ข้อสอบปรนัยจำนวน 298 ข้อที่สร้างจาก 12 ภาควิชาทางคลินิก ได้แก่ภาควิชาสูติศาสตร์-นรีเวชวิทยา ศัลยศาสตร์ กุมารเวชศาสตร์ อายุรศาสตร์ จิตเวชศาสตร์ เวชศาสตร์ป้องกันและสังคม โสต นาสิก ลาริงซ์วิทยา รังสีวิทยา วิสัญญีวิทยา นิติเวชศาสตร์ จักษุวิทยา และออร์โทปิดิกส์และเวชศาสตร์ฟื้นฟู ความเที่ยงของแบบสอบมีค่าเท่ากับ 0.85 เป็นข้อสอบที่ดีเยี่ยมจำนวน 21 ข้อ (ร้อยละ 7.05) เป็นข้อสอบที่ดีจำนวน 32 ข้อ (ร้อยละ 10.74) เป็นข้อสอบที่ควรได้รับการแก้ไขจำนวน 20 ข้อ (ร้อยละ 6.71) เป็นข้อสอบที่ไม่ดีจำนวน 196 ข้อ (ร้อยละ 65.77) และเป็นข้อสอบเชิงลบ จำนวน 29 ข้อ (ร้อยละ 9.73)

ข้อสอบที่ดีเยี่ยมและข้อสอบที่ดีเป็นข้อสอบที่มีประโยชน์มาก สมควรที่จะเก็บไว้ในคลังข้อสอบ ผู้ออกข้อสอบควรมีการตรวจสอบทั้งคำถามและตัวเลือก ผู้บริหารควรมีนโยบายในการวิเคราะห์ข้อสอบเป็นประจำทุกปี เพื่อเป็นการประกันคุณภาพของข้อสอบ

In 1979 the Faculty of Medicine, Chulalongkorn University developed the M.D. curriculum for serving national policy. At that time the Government desired that the medical school increase the numbers of medical graduates, and the time of the curriculum should be decreased. The curriculum used a 6 year course and graduates would have clinical competence sufficient for the medical council's criteria. For quality assurance, the comprehensive examination committee which had taken responsibility for constructing a test and managing the examination, constructed the Table of Specification and used it for constructing the test. The Table of Specification was Constructed based on the medical council's criteria.⁽¹⁾

The committee used the comprehensive examination to test sixth year medical students studying since 1976.⁽²⁾ At that time the students could receive a bachelors degree if they had a grade point average higher than or equal to 2.00, and had passed the comprehensive examination. The test consisted of 500 multiple choice question. After developed curriculum, if the sixth year medical students had a grade point average higher than or equal to

2.00 and had passed the comprehensive examination, they would receive a bachelors degree from the university and professional license from the medical council. Item Analysis techniques constitute some of the most valuable tools that a classroom teacher can apply in attempting to improve the quality of his tests. Item analyses are conducted for four general purposes: (1) to select the best available items for the final form of a test; (2) to identify any structural or content defects in any of the items; (3) to detect learning difficulties of the class as a whole or identifying general content areas or skills that need to be reviewed by the instructor; and (4) to identify for individual students areas of weakness which may be in need of remediation. There are three main elements involved in performing an item analysis. One is concerned with an examination of the difficulty level of the items. Another element involves determining the discriminating power of each item. The third element would involve an examination of the effectiveness of the distractors (alternative answers).⁽³⁾ In the past there was only one paper which studied the quality of the comprehensive examination MCQ test and that was for

Table 1. Quality of the 1993 comprehensive medical examination MCQ test. (n=298)

| Quality of question | Number of items | Percentage |
|---------------------|-----------------|---------------|
| Excellent | 21 | 7.50 |
| Good | 32 | 10.74 |
| Revise | 20 | 6.71 |
| Poor | 196 | 65.77 |
| Negative | 29 | 9.73 |
| Total | 298 | 100.00 |

the academic year 1979.⁽⁴⁾ The authors of this study decided to determine the quality of the MCQ test which was used in the academic year 1993.

Objectives

1. To study the 1993 comprehensive examination MCQ test in terms of item-difficulty index, item-discrimination index and reliability of the test.
2. To identify numbers of excellent, good, revise, poor and negative question by using the standard criteria of the World Health Organization.⁽⁵⁾

Definitions

1. Excellent question: a question having a discrimination index in the range 0.35 and over and the difficulty index lies between 0.30 and 0.70.
2. Good question: a question having a discrimination index in the range 0.25 to 0.34 and the difficulty index lies between 0.30 and 0.70.
3. Revise question: a question having a discrimination index in the range 0.25 and over and the

difficulty index does not lie between 0.30 and 0.70.

4. Poor question: a question having a discrimination index under 0.25.

5. Negative question: a question having a negative discrimination index. This question scored correct more frequently by 'bad' than 'good' students.

Materials

1. One microcomputer.
2. One printer.
3. One optical reader machine.
4. Software
 - 4.1 TOOLS: optical reader machine software.
 - 4.2 CTIA: item analysis software.
 - 4.3 QEdit: word processing software.
5. Computer MCQ answer sheets.

Methods

1. Use the optical reader machine to scan the answer sheets for obtaining raw data.
2. Use word processing to prepare the raw data for the CTIA format.

3. Run the item analysis program to identify the numbers of excellent, good, revise, poor, and negative questions and to calculate as percentages.

Results

1. In the 1993 comprehensive medical examination MCQ test there were 298 multiple choice questions. The reliability of this test was determined to be 0.85. When using WHO's criteria to identify the questions it was determined that there were 21 excellent questions (7.05%), 32 good questions (10.74%), 20 revise questions (6.71%), 196 poor questions (65.77%) and 29 negative questions (9.73) [Table 1].

2. The 1993 comprehensive medical examination MCQ was written by instructors from 12 clinical departments. It contained 298 questions, 46 from the Department of Obstetrics and Gynecology, 61 from the Department of Surgery, 44 from the Department of Pediatrics, 63 from the Department of Medicine, 8 from the Department of Psychiatry, 24 from the Department of Preventive and Social Medicine, 6 from the Department of Oto-laryngology, 8 from the Department of Radiology, 11 from the Department of Anesthesiology, 5 from the Department of Forensic Medicine, 5 from the Department of Ophthalmology, and 17 from the Department of Orthopedic Surgery and Rehabilitation medicine.

Table 2. Quality of the questions From the 12 departments.

| Department | Excellent | Good | Revise | Poor | Negative | Total |
|---------------------|----------------------------|-----------------------------|----------------------------|------------------------------|----------------------------|------------------------------|
| OB-GYN | 2 (4.35) | 1(2.17) | 5(10.87) | 54(73.91) | 4 (8.7) | 46 |
| Surgery | 4 (6.56) | 6(9.84) | 4(6.56) | 41(67.2) | 6(9.84) | 61 |
| Pediatrics | 1 (2.27) | 4(9.09) | 1(2.27) | 34(77.28) | 4(9.09) | 44 |
| Medicine | 6 (9.52) | 8(12.7) | 5(7.94) | 37(58.73) | 7(11.11) | 63 |
| Psychiatry | - | - | - | 8(100.0) | - | 8 |
| Preventive | 3 (12.5) | 5(20.83) | 1(4.17) | 11(45.83) | 4(16.67) | 24 |
| Oto-laryngo. | - | 2(33.33) | - | 4(66.67) | - | 6 |
| Radiology | 1 (12.5) | 2(25.0) | 1(12.5) | 4(50.0) | - | 8 |
| Anesthesio. | 2(18.18) | 1(9.09) | 1(9.09) | 7(63.64) | - | 11 |
| Forensic Med. | - | - | 3(60.0) | 2(40.0) | - | 5 |
| Ophthalmology | 1(20.0) | - | 4(80.0) | - | - | 5 |
| Orthopedic. | 1(5.88) | 3(17.65) | 2(11.76) | 9(52.95) | 2(11.76) | 17 |
| Total (%) | 21 (7.05) | 32 (10.74) | 20 (6.71) | 196 (65.77) | 29 (9.73) | 298 (100.0) |

Table 3. Range and mean of the difficulty index and discrimination index of 12 subjects in the 1993 comprehensive medical examination MCQ test.

| Subject | Difficulty Index | | Discrimination Index | |
|---------------|------------------|------|----------------------|------|
| | range | mean | range | mean |
| OB-GYN | 0.000 - 0.993 | .788 | -0.128 - 0.436 | .110 |
| Surgery | 0.081 - 0.959 | .647 | -0.077 - 0.487 | .152 |
| Pediatrics | 0.074 - 0.993 | .660 | -0.077 - 0.385 | .117 |
| Medicine | 0.000 - 0.986 | .541 | -0.205 - 0.615 | .164 |
| Psychiatry | 0.203 - 0.932 | .585 | -0.000 - 0.231 | .080 |
| Preventive | 0.081 - 0.959 | .520 | -0.154 - 0.385 | .171 |
| Oto-laryngo. | 0.338 - 0.885 | .579 | -0.026 - 0.308 | .192 |
| Radiology | 0.318 - 0.986 | .620 | -0.051 - 0.359 | .228 |
| Anesthesio. | 0.189 - 0.730 | .453 | -0.051 - 0.487 | .198 |
| Forensic Med. | 0.757 - 0.953 | .853 | -0.077 - 0.179 | .046 |
| Ophthalmology | 0.642 - 0.980 | .820 | 0.000 - 0.359 | .169 |
| Orthopedic. | 0.108 - 0.953 | .582 | -0.128 - 0.487 | .184 |

The questions from 6 departments (OB-GYN, Surgery, Pediatrics, Medicine, Preventive Med., Orthopedic.) had excellent, good, revise, poor, negative questions. There were 5 departments which had no negative questions. For the department of Forensic medicines 5 questions 3 (60%) were poor and 2 (40%) were negative questions. The department of Psychiatry had 8 questions and all of them (100%) were determined to be poor questions [Table 2].

3. Regarding the difficulty index, three departments (OB-GYN, Forensic Medicine, Ophthalmology) had easy questions ($p \geq 0.78$ by average). The remaining nine departments (Surgery, Pediatrics, Medicine, Psychiatry, Preventive Med., ENT, Radiology, Anesthesiology, Orthopedic. Med.) had fair questions ($p = 0.45-0.66$ by average). The questions of all departments had an average discrimination index of less than 0.20 except Department of Radiology which had a discrimination index averaging 0.23 [Table 3].

Discussion

Out of 298 questions in the MCQ test, there were 53 (17.79%) high quality questions (excellent + good question). These would be most useful for retaining in a question bank. (5) Revise questions, which were questions with high discrimination index, could help teachers to determine who were the best students. A problem of the revise questions was their difficulty index not being between 0.30 and 0.70. This criteria was based on World Health Organization suggestion (6) but, in general, a question with a difficulty index between 0.20 and 0.80 was acceptable. (7,8) There were 196 poor questions (65.77%). These were mostly questions that were too easy or too difficult. In educational psychology, the easy questions would be useful if used in the first part of tests as they could motivate students to do the whole test. Another reason was the poor quality of distractors. A good way to derive distractors is to determine the common errors students make, and the common misconceptions they hold. The crucial point of poor questions was that if they had content validity, they were valuable and the instructor concerned could accept them. In this study, all questions from Department of Psychiatry were poor. The instructor should seek ways to improve them. The negative questions must be a crucial point for which all instructors should be concerned. The questions were correctly answered by more poor students than good students. These questions should be examined carefully to clarify why good students did not choosing the correct answer. (9) However, this MCQ was determined to be a good test as. Hubbard and Clemans, (10) Schumacher, (11) Cox and Ewan (12) suggested that a good test should have reliability 0.70 or over.

The first paper analysing the comprehensive medical examination MCQ test was published in 1981. (4) Our study is the second such paper. The construction and analysis of the MCQ test was hard work and required considerable time. The Comprehensive Medical Examination Committee might check the quality of the questions

in terms of validity, objectivity, and comprehensiveness item by item. At the last meeting of the 1993 Comprehensive Medical Examination Committee, they summarized the problems and provided recommendation. (13) The Committee members felt tired and bored because there was little incentive for them. The MCQs from some departments were not complete and were sent to the Committee late. To solve these problems, the faculty should have 1) the incentive system, and accept this work to be part of the instructor's work load. 2) permanent officers who take full responsibility for the comprehensive examination. 3) a policy for analysing every MCQ test which has been used in the comprehensive medical examination and kept in the item bank and 4) the term of the Comprehensive Medical Examination Committee should be two years and half of members should be changed every year.

Summary

This descriptive research aimed to study the quality of the MCQ test. The 1993 Comprehensive, Medical Examination of the Faculty of Medicine, Chulalongkorn University in terms of item-difficulty index, item-discrimination index, reliability of the test and to determine the number of excellent, good, revise, poor and negative questions by using the standard criteria of the World Health Organization. The MCQ test was analysed by use of the CTIA item analysis program. The test consisted of 298 questions from 12 clinical departments: Obstetrics and Gynecology, Surgery, Pediatrics, Medicine, Psychiatry, Preventive and Social Medicine, Oto-laryngology, Radiology, Anesthesiology, Forensic Medicine, Ophthalmology and Orthopedic Surgery and Rehabilitation Medicine. The reliability of test was determined to be 0.85. There were 21 excellent questions (7.05%), 32 good questions (10.74%), 20 revise questions (6.71%), 196 poor questions (65.77%) and 29 negative questions (9.73%). Both excellent and good questions are the most useful and should be kept in the question bank. Instructors who wrote the questions should completely check the questions, stem and distractors. The program director should have a policy of analysing the test every year for the sake of quality control.

References

1. Faculty of Medicine, Chulalongkorn University. Table of Specification for Comprehensive Examination. Bangkok: Division of Academic Affairs, 1993.1
2. Phulklongtan M, Jaroongdaechakul M, Limpapayom K. An analysis of comprehensive examination items of a cademic year 1979. Chula Med J 1981 Sep; 25(5):1036
3. Payne DA. The specification and measurement of learning outcomes. Toronto: Xerox College, Publishing, 1968:145
4. Phulklongtan M, Jaroongdaechakul M, Limpapayom K. An analysis of comprehensive examination items of academic year 1979. Chula Med J 1981

- Sep; 25(5):1035-40
5. Guilbert JJ. Educational Handbook for Health Personnel. Revised ed. Geneva: World Health Organization, 1987: 4.68-4.69
 6. Guilbert JJ. Educational Handbook for Health Personnel. Revised ed. Geneva: World Health Organization, 1987: 4.69
 7. Kasemsunti C, Wachirapornthip A. Item analysis. Siriraj Hosp Gaz 1975 Aug; 27(8):1181-96
 8. Sukamolson S. Classical Test Item Analysis and Grading manual (V.6.30). Bangkok: Chulalongkorn University Language Institute, 1992:36
 9. Cox K, Ewan CE. The Medical Teacher. 2nd.ed. London:Churchill Livingstone, 1988:164
 10. Hubbard JP, Clemans WV. Multiple-Choice Examinations in Medicine: A Guide for Examiner and Examinee. Philadelphia: Lea & Febiger, 1961:71
 11. Schumacher CF. Scoring and analysis. In: Hubbard JP, ed. Measuring Medical Education. Philadelphia: Lea & Febiger, 1971.
 12. Cox K, Ewan CE. The Medical Teacher. 2nd.ed. London: Churchill Livingstone, 1988: 163
 13. Faculty of Medicine, Chulalongkorn University. Final Report of the 1993 Comprehensive Examination Committee on April 29, 1994. Bangkok: Division of Academic Affairs, 1994.