Psychometric analysis pattern of Pediatrics course in college of Medicine. KKU ORAL

Dr. Alam Eldin Musa Mustafa Musa, Assistant Professor Pediatric department college of Medicine , KKU, alammusa68@gmail.com

Abstract - Assessment is a basic component of training and assumes a significant job in offering inputs to partners. Understudies appraisal and assessment are an indispensable piece of the instructing and learning process (Popham, 2002). The educator's information in appraisal and assessment is anything but a static procedure but instead a mind boggling, dynamic, and continuous movement. These Authors additionally contend that this kind of information creates along a worldly continuum; educators ordinarily utilize their previous appraisal experience to design their present evaluation rehearses. Among the various kinds of understudies learning accomplishments and progress the numerous decision questions are all inclusive the most used (Swanson et al., 2004)[1-2]

I. INTRODUCTION:

Thing examination is a procedure which looks at understudy reactions to singular test things (questions) so as to survey the nature of those things and of the test all in all. Thing investigation is particularly important in improving things which will be utilized again in later tests, however it can likewise be utilized to kill vague or deluding things in a solitary test organization. The procedure of psychometric investigation help in deciding the trouble list (DI), separation record (DIS), dis-tractor file (DE) and dependability of the tests.

Reliability coefficients hypothetically extend in an incentive from zero (no dependability) to 1.00 (immaculate unwavering quality). Trouble list (p-esteem), likewise called ease record, depicts the level of understudies who effectively addressed the thing. It ranges from 0 - 100%. The higher the rate, the simpler the thing. The suggested scope of trouble is from 30 - 70%. Things having p-values underneath 30% or more 70% are viewed as troublesome and simple things individually.[3-4]

Difficulty Index:

Difficulty index additionally called ease list, portrays the level of understudies who accurately addressed the thing. It ranges from 0 - 100%. The higher the rate, the simpler the thing. The suggested scope of trouble is from 30 - 70%. Things having p-values beneath 30% or more 70% are viewed as troublesome and simple things individually.[5-6]

Distractor Index:

Investigation of distractor is another significant piece of thing examination. The distractors are significant segments of a thing, as they show a connection between the all out grade and the distractor picked by the understudy. execution relies on how the distractors are planned. Distractor productivity is one such device that tells whether the thing was very much developed or neglected to play out its motivation. Any distractor that has been chosen by under 5% of the understudies is viewed as a non-working distractor (NF-D) Ideally, low-scoring understudies, who have not aced the subject, ought to pick the distractors all the more regularly, though, high scorers should dispose of them all the more every now and again while picking the right alternative. By examining the distractors, it gets simpler to distinguish their blunders, with the goal that they might be reconsidered, supplanted, or expelled.[1-8]

Reliability: (DE)

Reliability coefficients hypothetically extend in an incentive from zero (no dependability) to 1.00 (flawless unwavering quality). 0 .90 or more Excellent unwavering quality; at the degree of the best state administered test 0.80 – 0.90 Very useful for a homeroom test 0.70 – 0 .80 Good for a study hall test; in the scope of most. There are most likely a couple of things which could be improved. 0 .60 – 0.70 Some what low. This test should be enhanced by different measures (e.g., more tests) to decide grades. There are most likely a few things which could be improved. 0 .50 – 0.60 Suggests requirement for update of test, except if it is very short (ten or less things). The test certainly should be enhanced by different measures (e.g., more tests) for grading.0.50 or underneath Questionable unwavering quality. This test ought not contribute vigorously to the course evaluation, and it needs amendment.[1-8]

Discrimination Index:

The Discrimination Index (DIS), otherwise called the point biserial connection, which distinguishes segregation between understudies with various degrees of accomplishment; and distractor proficiency (DE), which demonstrates whether the distractors in the thing are all around picked or have neglected to occupy understudies from choosing the right answer. A perfect thing ought to have a DI > 0.2[1-8]

Objective:
Compare the progress of pediatrics 2 item analysis predictors from 2017-2019 , further to discuss the possible ways to improve the standards of multiple choice questions MCQ's)

II. METHODOLOGY:

It is a Retrospective cross sectional study, data of all final exams (Pedia course in MBBS program) were collected from examination department then after that enter into the SPSS software for calculations and finding the values of the item analyses indexes. Data for three years was analyzed

Difficulty Index (DI):

 Students with correct answer/ Total students) x 100

Discrimination Index (DIS):
DIS　　　　　　　　　　　= (H　　　　　　　　　　　　　　－　　　　　　　　　　　　　　L)
　　27% of Total
H = number of correct answers from top 27% of students
L = number of corrects answers from bottom 27% of students

Reliability: (Kuder and Richardson Formula 20) (DE)

$$\rho_{KR20} = \frac{k}{k-1}\left(1 - \frac{\sum_{j=1}^{k} p_j q_j}{\sigma^2}\right)$$

Where

k = number of questions
p_j = number of people in the sample who answered question j correctly
q_j = number of people in the sample who didn't answer question j correctly

σ^2 = variance of the total scores of all the people taking the test = VARP(R1) where R1 = array containing the total scores of all the people taking the test.
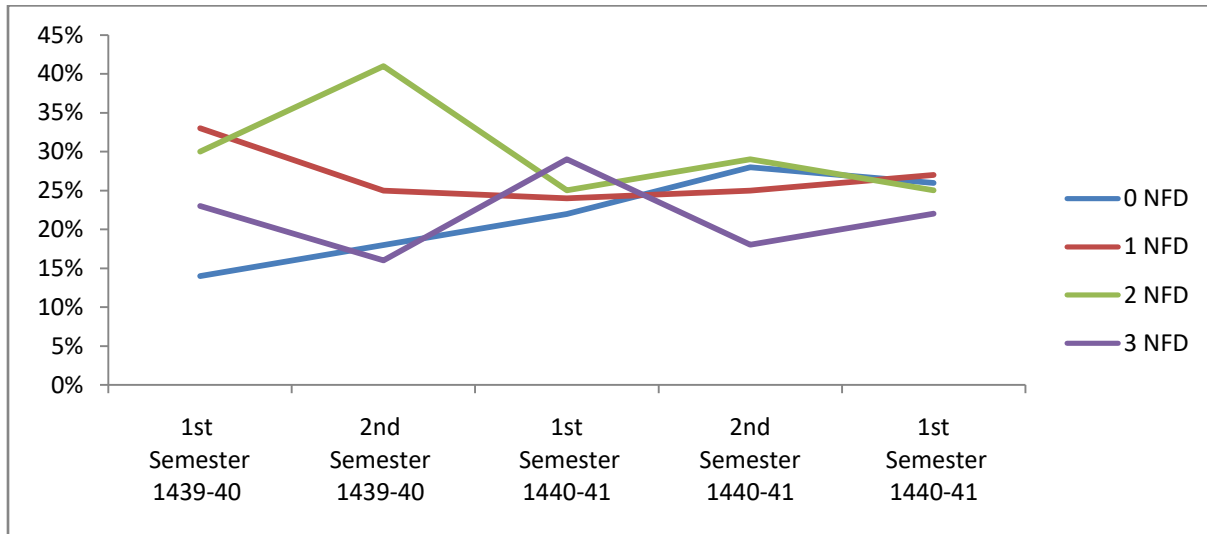
III. RESULTS:

Table 1:

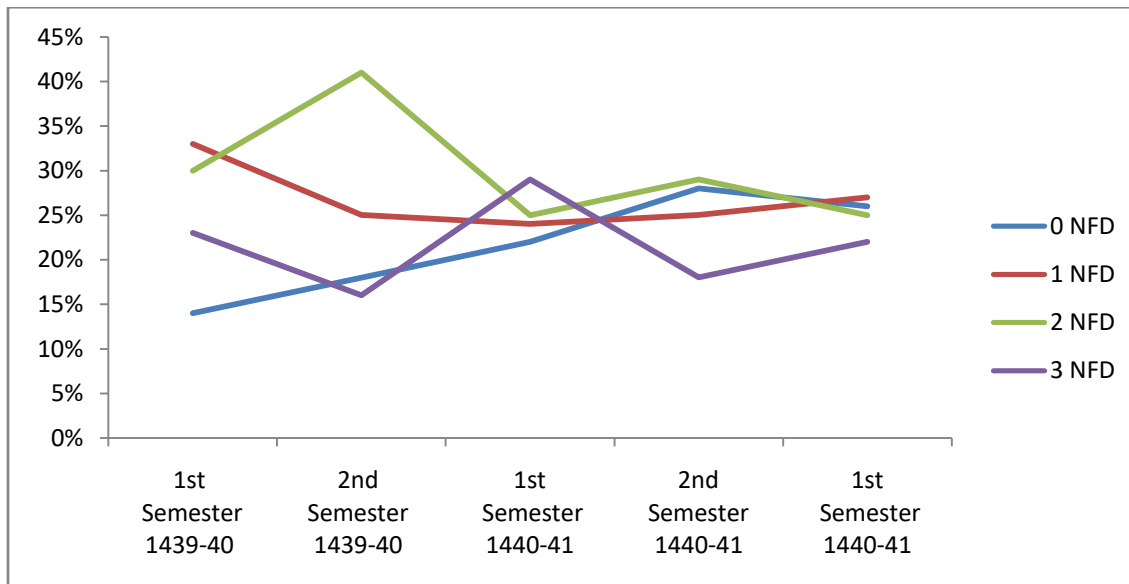|	1439-40		1440-41		1441-42	P- value
	1st Semester 1439-40	2nd Semester 1439-40	1st Semester 1440-41	2nd Semester 1440-41	1st Semester 1440-41	
Difficulty Index						
Difficult questions	10%	15%	20%	35%	8%	P<0.05
Easy questions	60%	35%	35%	24%	37%	P<0.05
Moderate questions	30%	50%	45%	51%	55%	P<0.05
Reliability (K-20)						
	0.54	0.68	0.74	0.54	0.74	N.S
Discrimination Index						
Negative question	15%	5%	8%	7%	5%	P<0.05
Above 0.2 value	53%	55%	60%	62%	63%	N.S
Distractor Index						
0 NFD	14%	18%	22%	28%	26%	P<0.05
1 NFD	33%	25%	24%	25%	27%	N.S
2 NFD	30%	41%	25%	29%	25%	N.S
3 NFD	23%	16%	29%	18%	22%	N.S

A total of 150 MCQ items were prepared for final exams/ summative assessment (during 2017-2019 first semester) by the faculty members of COM, KKU Pedia 02 course, each items have one correct question and three distractors. The instructors tried to prepare the MCQ items as per guidelines of KKU. Likewise, the quality (i.e., good construction) of the constructed MCQ items wasassessed with the help of discrimination index (DI) values. The MCQ items' DI values werefound proportional of good/poor constructed questions ratio (%), and we have observed the increase in the values of above0.2 values significantly. Further analysis of the MCQ items revealed a significant increase in the questions with functional distractors.
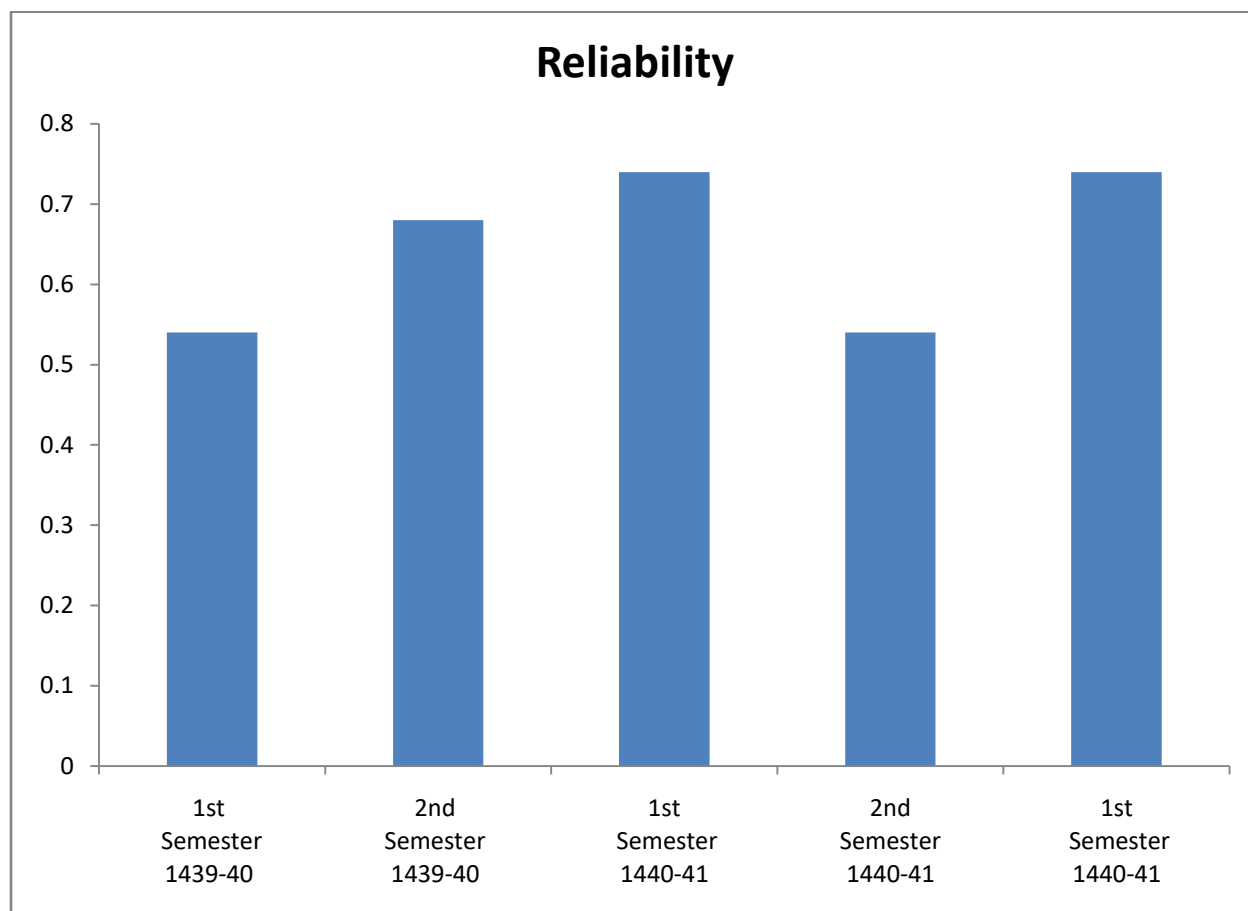
0 NFD are increasing

Figure2:

Difficulty Index:

Moderate questions are increasing…

Figure 3:

Reliability

Bar chart showing reliability values for:
- 1st Semester 1439-40: ~0.54
- 2nd Semester 1439-40: ~0.68
- 1st Semester 1440-41: ~0.74
- 2nd Semester 1440-41: ~0.54
- 1st Semester 1440-41: ~0.74

IV. DISCUSSION:

Different decision questions (MCQs) are one of the well known and acknowledged methods for assessment in clinical instruction. MCQ test things are worthwhile as they can cover more extensive area of exercises and examine enormous quantities of understudies in lesser time all the while. The tests can be utilized for the two ideal models of appraisal (developmental and summative). Schools are consolidating MCQs tests in their assessments as there is raising pattern of receiving MCQs for postgraduate clinical placement tests. Its acknowledgment depends on its objectivity, possibility, high inner consistency and precision. [9-11]

Different examinations have audited the MCQs test things utilized for the understudies' evaluation before and found a high level of things composing imperfections (IWFs) and non-utilitarian distractors (NFD) . Because of the nearness of previously mentioned and other dangerous components in the MCQs test things, committed MCQ test things composing preparing by the clinical master was suggested by the above creators. Having confined MCQs as indicated by the rules, itis essential to examine the nature of these things and whether these can separate high and low capacity understudies. Steinert et al. in their methodical audit in wellbeing callings instruction announced that staff preparing was seen as related decidedly with showing viability both quick and long haul. FDPs hence is pivotal for advancement of substantial and dependable appraisal material. A great MCQ is one in which each distractor is chosen by probably a few understudies who don't have the foggiest idea about the substance tried by the inquiry. Finding of the level of information and the idea of an examinee's misconception can be made by evaluating the usefulness of distractors (Nitko 2004; Popham 2000). [12-17].Results from this examination show that the relative extents of simple, troublesome, review, non-segregating and non-practical things were reliably expanding.

The chance to examine and fundamentally audit questions empowered the employees to show their insight into understudy learning in their order. This was obvious in the sort of inquiries picked to survey learning results (e.g.application versus review) and furthermore in the choice and method of reasoning for the utilization of distractors in the MCQ.A great MCQ is one in which each distractor is chosen by in any event somestudents who don't have the foggiest idea about the substance tried by the inquiry. Conclusion of the level of information and the idea of an examinee's misconception can be made by surveying the usefulness of distractors (Nitko 2004; Popham 2000)[14]. Given that evaluation drives learning, a distractor that neglects to draw in any examinees is broken, doesn't aid the estimating of instructive results, adds nothing to the thing or the test (psychometrically) and has negative effect upon students.

V. REFERENCES:

1. Swanson JW, Swartz MS, Elbogen EB. Effectiveness of atypical antipsychotic medications in reducing violent behavior among persons with schizophrenia in community-based treatment. Schizophr Bull. 2004;30:3-20.
2. Link BG, Stueve A. Psychotic symptoms and the violent/illegal behavior of mental patients compared to community controls. In: Monahan J, Steadman HJ, eds.
3. Violence and Mental Disorders: Developments in Risk Assessment. Chicago, Ill:
4. University of Chicago Press; 1994:137-160

5. Hingorjo MR, Jaleel F. Analysis of one-best MCQs: The difficulty index, discrimination index and distractor efficiency. J Pak Med Assoc 2012; 62:142–7. 7. Tarrant M, Ware J, Mohammed AM.
6. An assessment of functioning and non-functioning distractors in multiple-choice questions: A descriptive analysis. BMC Med Educ 2009; 9:40. doi: 10.1186/1472-6920-9-40
7. Tarrant M, Ware J, Mohammed AM. An assessment of functioning and non-functioning distracters in multiple-choice questions: a descriptive analysis. BMC Med Educ 2009; 9: 40.
8. Gronlund NE, Linn RL. Measurement and evaluation in teaching. 6th ed. New York: Macmillan publishing Co; 1990.
9. Chandratilake M, Davis M, Ponnamperuma G. Evaluating and designing assessments for medical education: the utility formula. Intern J Med Edu. 2009;1:1-7.
10. Hingorjo MR, Jaleel F. Analysis of one-best MCQs: the difficulty index, discrimination index and distracter efficiency. J Pak Med Assoc. 2012; 62:142-6.
11. McCourbie P. Improving the fairness of multiple-choice questions: a literature review. Med Teach 2004; 26: 709–712. doi: 10.1080/01421590400013495
12. Scorepak®: Item analysis. [Last citied on 2017 Apr 30]. Available from: www.washington.edu./oea/score1/htm .
13. Matlock-Hetzel S. Presented at annual meeting of the Southwest Educational Research Association, Austin, January; 1997. [Last cited on 2017 May 2]. Basic concept in item and test analysis. Available from: www.ericae.net/ft/tamu/espy.htm .
14. Popham, WJ. (2004). Why assessment illiteracy is professional suicide. Educational leadership: journal of the Department of Supervision and Curriculum Development, N.E.A. 62. 82-83.
15. Yvonne Steinert, Karen Mann, Brownell Anderson, Bonnie Maureen Barnett, Angel Centeno, Laura Naismith, David Prideaux, John Spencer, Ellen Tullo, Thomas Viggiano, Helena Ward & Diana Dolmans (2016) A systematic review of faculty development initiatives designed to enhance teaching effectiveness: A 10-year update: BEME Guide No. 40, Medical Teacher, 38:8, 769-786, DOI: 10.1080/0142159X.2016.1181851
16. Alford DP, Richardson JM, Chapman SE, Dubé CE, Schadt RW, Saitz R. 2008. A web-based Alcohol Clinical Training (ACT) curriculum: is in-person faculty development necessary to affect teaching? BMC Med Educ 8(1):11. [Google Scholar]
17. Amin Z, HoonEng K, Gwee M, ChayHoon T, Dow Rhoon K. 2006. Addressing the needs and priorities of medical teachers through a collaborative intensive faculty development programme. Med Teach 28(1):85–88. [Taylor & Francis Online], [Web of Science ®], [Google Scholar]