

Effect of surprise test and instruction for negative marking on item analysis in Pharmacology

Nandal Dattatraya Hanumantrao¹, Narwane Sandeep Prakash^{2*}

¹Professor and Head, ²Associate Professor, ^{1,2}Dept. of Pharmacology, ¹Birat Medical College Teaching Hospital, Tankisinuwari, Nepal, ²Rural Medical College, Loni, Maharashtra, India

*Corresponding Author: Narwane Sandeep Prakash

Email: drsandeepnarwane1984@gmail.com

Abstract

Introduction: The effect of instructions of negative marking on item analysis during an MCQ examination was studied.

Material and Methods: It was an observational type of longitudinal study conducted in MBBS students of second year. The students were given the same MCQ test twice with surprise. There were no instructions regarding negative marking during the first test (T1), whereas instructions were given for the second test (T2). The Raw score, Negative score and Corrected score along with Difficulty and Discrimination index of each MCQ was calculated.

Results: There was no statistically significant between the indices when the values of T1 and T2 were compared. There was statistically significant inverse correlation found between the difficulty index and the number of students who did not attempt the respective MCQ (Spearman correlation coefficient= -0.7818).

Conclusion: The Difficulty Index and Discrimination Index do not change with Conventional test and test with instructions of negative marking. Thus, these indices form an important tool for MCQs irrespective of the conventional and negative marking.

Keywords: Difficulty index, Discrimination index, Multiple choice questions, Negative marking, Surprise test.

Introduction

MCQs (Multiple Choice Questions) are the integral part of assessment of medical students in formative and summative assessment. A good quality MCQ assesses the knowledge and differentiates various abilities of a student.¹ MCQ consists of direction to students, a stem and alternatives. Usually an MCQ has multiple choices (generally four), of which one is correct while the remaining three are known as distracters.² Item analysis evaluates the reply of students to each MCQ and suggests its quality and scope of perfection. The indices, namely difficulty Discrimination indices are the important tools of item analysis.³

In a study by Narwane et al,⁴ it was observed that there was a negative impact on the performance of students, when the instructions of negative marking were given. Therefore it is also likely that the parameters of item analysis would also be affected. The literature on the effect of instructions of negative marking on the item analysis is scarce. Therefore, the present study therefore was planned to see the effect of instructions of negative marking on item analysis during an MCQ examination was studied.

Materials and Methods

It was an observational type of longitudinal study conducted in MBBS students of second year. The students were given the same 20 MCQ test twice with surprise. There were no instructions regarding negative marking during the first test (T1), whereas instructions were given for the second test (T2). The Raw score, Negative score and Corrected score along with Difficulty and Discrimination index of each MCQ was calculated. The students used their own alpha-numerical code as described by Narwane et al.⁴ All the coded papers were assessed by a teacher unaware of the

identity of students and the purpose of this educational study.

After the completion of assessment, difficulty and discrimination index during T1 and T2 for each MCQ were calculated using Raw score, Corrected score and Negative score. The negative score was calculated by the addition of one mark and subtraction of 1 mark for each right and wrong answer, respectively. Marks were neither subtracted nor added for a non-attempted question. The corrected score was calculated by the following formula

$$\text{Corrected score} = \frac{\text{Raw score} - \text{Number of wrong answers}}{(n - 1)}$$

Where n= Number of options for MCQ. Marks were neither subtracted nor added for a non-attempted question³.

Formula for difficulty index (DI)

$$\text{Difficulty index} = \frac{H + L}{N} \times 100$$

Formula for discrimination index (DCNI)

$$\text{Discrimination Index} = 2 \times \frac{H - L}{N}$$

Here, the students were ranked in a descending manner as per their overall performance. The upper 1/3rd and the lower 1/3rd formed the High score group and Low score group, respectively. H (among High score group) and L (among Low score group) are the number of students who answered the respective MCQ correctly. N is the number of students in high and low score group.

Results

Sixty seven students participated in the study.

Table 1: DI and DCNI values using Raw, Negative and Corrected score

| Q No | RS of T1 | RS of T2 | NS of T1 | NS of T2 | CS of T1 | CS of T2 |
|------|-------------|---------------------------|-------------|---------------------------|-------------|---------------------------|
| DI | 62.27±19.01 | 58.98±23.43 ^{NS} | 62.16±18.89 | 59.32±22.55 ^{NS} | 62.27±19.28 | 59.20±22.63 ^{NS} |
| DCNI | 0.35±0.15 | 0.34±0.22 ^{NS} | 0.34±0.16 | 0.30±0.18 ^{NS} | 0.35±0.16 | 0.33±0.17 ^{NS} |

^{NS}P>0.05 VS respective Raw/Negative/Corrected score of T1 (Paired t test)

DI-Difficulty index, DCNI-Discrimination index, RS-Raw Score, NS-Negative Score, CS-Corrected Score

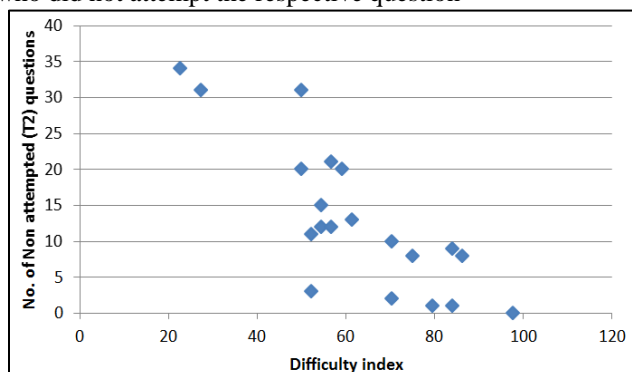
On comparing the respective DI calculated for Raw, Negative and Corrected score of T1 with that of T2, no significant difference was found (Table 1). Similar results were seen with the DCNI (Table 1).

Table 2: Correlation of difficulty index verses number students who did not attempt the respective question

| Q. No | DI | Number of Students NA (T2) | Number of Students NA (T1) |
|-------|----------|----------------------------|----------------------------|
| 16 | 22.72727 | 34 | 2 |
| 13 | 27.27273 | 31 | 3 |
| 12 | 50 | 31 | 3 |
| 10 | 50 | 20 | 2 |
| 8 | 52.27273 | 3 | 0 |
| 20 | 52.27273 | 11 | 2 |
| 3 | 54.54545 | 12 | 1 |
| 14 | 54.54545 | 15 | 1 |
| 2 | 56.81818 | 12 | 0 |
| 7 | 56.81818 | 21 | 1 |
| 6 | 59.09091 | 20 | 1 |
| 11 | 61.36364 | 13 | 0 |
| 1 | 70.45455 | 10 | 0 |
| 5 | 70.45455 | 2 | 0 |
| 9 | 75 | 8 | 0 |
| 17 | 79.54545 | 1 | 0 |
| 4 | 84.09091 | 9 | 3 |
| 19 | 84.09091 | 1 | 0 |
| 15 | 86.36364 | 8 | 0 |
| 18 | 97.72727 | 0 | 0 |
| | Total | 262 | 19 |

P<0.0001 (Pearsons Correlation test, r = -0.8154), NA- Non Attempting.

Fig. 1: Correlation of difficulty index and number students who did not attempt the respective question



As shown in Table no. 2 and Figure no. 1, there was statistically significant negative correlation between these parameters (P<0.0001, r = -0.8154, Pearsons correlation test).

Discussion

Allocating marks to right answer (Raw score) is the most commonly favored methods of assessment of multiple choice questions^{5,6}. To avoid guessing of answers, negative scoring is used, where negative marks are awarded for an attempted question with wrong answer⁷. It has been observed that there was increase in total score in students who guess answers.⁹⁻¹¹ Guessing also lowers the validity and reliability by introducing a random factor into test scores.^{12,13} Test designers cannot distinguish between correct answers attributable to knowledge mastery as compared to those based on a guess.¹⁴

However, it also discourages students from answering questions where they are less confident. The corrected score method takes into consideration the number of options provided in the MCQs. After excluding the right alternative, the student has probability of being wrong becomes 1/(n-1), for instance, 33% for an MCQ with four alternatives. Therefore, the marks subtracted for each wrong answer is 1/3 instead of 1, for opting one of wrong alternatives.

As shown in Table no. 3, on comparing the Difficulty indices for Raw, Negative and Corrected score of T1 with

that of T2, there was no statistical significance found (Paired t test). Similar results were seen on with Discrimination index (Paired t test). Hence these indices do not change irrespective of the system of marking as well as instruction for negative marking that were used for evaluation of the MCQs. There were no studies found in literature that studied and published regarding the differences in scoring system with respect to surprise test.

The Difficulty indices of each MCQ were correlated with the Number of times it was not attempted (Table 4, Fig. 1). There was highly significant negative correlation ($P < 0.0001$, Pearsons Correlation test, $r = -0.8154$) between the number of times the MCQs were not attempted in the T1 and T2. Thus, the probability of not attempting a question increases with increase in the level of difficulty of the question.

Conclusion

The Difficulty and Discrimination indices are valid parameters irrespective of the type of scoring system used for MCQs. The number of non attempted questions increase as the difficulty index decreases for a given multiple choice question.

Source of funding

None.

Conflict of interest

None.

References

- 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 J Comm Med* 2014;39(1):17-20.
- Kehoe J. Writing multiple-choice test items. *Pract Assess, Res Eval* 2008;4(9)
- Guilbert, Jean-Jacques & World Health Organization. Educational handbook for health personnel / J.-J. Guilbert, 6th ed. rev. and updated 1998. WHO. <https://apps.who.int/iris/handle/10665/42118> (accessed on 8th July 2019)
- Narwane SP, Nandal DH, Pawade RB, Kunkulol RR, Patil GD. (2019). Lessons learned from performance of students of Pharmacology in self coded surprise test with negative marking. *Int J Clin Biomed Res* 2019;5(3):24
- Meyer BY, Meyer Y & Flascher OM. Prospect theory analysis of guessing in multiple choice tests. *J Behav Decis Mak* 2002;15:313-27.
- Kurz TB. A review of scoring algorithms for multiple-choice tests. Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio, TX. 1999.
- Choppin BH. Correction for guessing. In J. P. Keeves (Ed.), Educational research, methodology, and measurement: an international handbook Oxford: Pergamon Press. 1988;384-6.
- Budescu D & Bar-Hillel M. To guess or not to guess: a decision-theoretic view of formula scoring. *J Educ Meas* 1993;30(4):277-91.
- Frary RB. The effect of misinformation, partial information, and guessing on expected multiple-choice test item scores. *Appl Psychological Meas* 1980;4:79-90.
- Kubinger KD, Holocher-Ertl S, Reif M, Hohensinn C & Frebort M. On minimizing guessing effects on multiple-choice items: superiority of a two solutions and three distractors item format to a one solution and five distractors item format. *Int J Sel Assess* 2010;18(1):111-5.
- Meyer B. Quantifying the effects of chance in multiple choice and true/false tests: question selection and guessing of answers. *Assess Eval Higher Educ* 2002;26(1):41-50.
- Kubinger. Correcting for guessing increases validity in multiple-choice examinations in an oral and maxillofacial pathology course. *J Dent Educ* 2010;70(4):378-86.
- Bar-Hillel M, Budescu D and Attali Y. Scoring and keying multiple choice test: a case study in irrationality. *Mind & Soc* 2005;4:3-12.

How to cite: Nandal DH, Narwane SP. Effect of surprise test and instruction for negative marking on item analysis in Pharmacology. *IP Int J Comprehensive Adv Pharmacol* 2019;5(1):19-21.