

Modified Negative Marking Schemes in Multiple Choice Questions in a Health Institution in Nigeria

Ernest Ndukaife Anyabolu^{1,2,*}, Innocent Chukwuemeka Okoye¹

¹Department of Medicine, Chukwuemeka Odumegwu Ojukwu University, Awka, Nigeria

²Department of Medicine, Imo State University, Orlu, Nigeria

Abstract Background and Objectives: The use of multiple choice questions (MCQs) is an objective method of assessment of knowledge. The true/false pattern of MCQs is often applied in medical schools. Negative marking for the true/false MCQs may influence the overall reliability, validity, and outcome of the assessment. This study was set out to evaluate aspects of negative marking schemes in MCQs in a medical school in Nigeria. **Methodology:** This was a cross sectional study involving a set of final year medical students. Multiple choice questions comprising of 100 true/false questions and 70 best of five (BOF) options were administered in 2 sessions, 7 days apart, and 3 marking schemes were used. Scheme A: “Informed negative marking but no negative marking” Scheme B: “Informed no negative marking and no negative marking”. Scheme C: “Informed negative marking and negative marking”. Number right scheme was used for the BOF. The results were compared between the 3 marking schemes. **Results:** The students’ mean score 46.1% for Scheme A was higher than 39.9% for Scheme C ($p < 0.001$), but lower than 57.7% for Scheme B, ($p = 0.028$). The differences in the positive scores, as well as the negative scores between the three schemes were not significant. ($df = 22$, $p = 0.288$) ($df = 20$, $p = 0.862$) There was a near perfect correlation between the mean score for Scheme A and that for Scheme C ($r = 0.97$ $p < 0.001$). However, there was a moderate correlation between the mean score for Scheme A and Scheme B. The mean score for Scheme C was a predictor of the mean score for Scheme A. **Conclusion:** The Marking Schemes for MCQs “Informed negative marking but no implementation of negative marking” and “Informed negative marking and implementation of negative marking” were closely related, and were more reliable and valid than the Marking Scheme “Informed no negative marking and no implementation of negative marking” for the true/false MCQs. Negative marking should be adopted in true/false options in MCQs in medical schools.

Keywords Multiple choice questions, Negative marking schemes, True or false questions, Informed negative marking, Informed no negative marking, Nigeria

1. Introduction

The use of multiple choice questions (MCQs) for the assessment of knowledge is common in higher institutions [1]. The essay forms of written examinations seem to be giving way for MCQs; some centers still use a combination of MCQs and essays [2].

Unlike essay forms, MCQs have been shown to be more objective in assessing written knowledge [3].

Studies have shown that MCQs could adopt negative or non-negative marking schemes [4, 5]. In a negative marking scheme, wrong answers are penalized with negative marks, unanswered questions zero, whereas in non-negative marking wrong answers and unanswered questions are given no marks, that is, zero [6].

Guessing by students in MCQs may reduce the reliability of the tests [6, 7]. However, studies have demonstrated that guessing tends to decline with heavier penalties when the negative marking is prominent, and increases with little penalties in negative marking and non-negative marking [6, 8, 9].

The concept of fair penalty in a negative marking scheme to discourage distracter has been put at less than one where the mark allotted to a stem is 1 [6, 9].

The reliability, validity and objectivity of MCQs tend to increase as the number of questions increases [6, 10, 11]. The tests efficiency might be influenced by the variability in the number of stems in each question as well as the marking scheme [6, 11, 12].

There is a paucity of studies on MCQs marking schemes in higher institutions in Nigeria. This has prompted us to undertake this study to evaluate negative marking schemes in MCQs in the assessment of knowledge in a higher institution in Nigeria.

* Corresponding author:

enhealer@yahoo.com (Ernest Ndukaife Anyabolu)

Published online at <http://journal.sapub.org/cmd>

Copyright © 2017 Scientific & Academic Publishing. All Rights Reserved

2. Materials and Methods

This was a cross-sectional study involving 34 final year medical students preparing for their final MBBS examinations. It was conducted in the Department of Medicine, Chukwuemeka Odumegwu Ojukwu University, Nigeria, in June and July 2017.

Multiple choice questions (MCQs) comprising of 100 true/false questions and 70 best of five (BOF) options were administered in 2 sessions, 7 days apart. For the BOF questions Number Right (NR) was used in which +1 mark was awarded for a correct answer, no mark for unanswered question or wrong answer. For the 100T/F questions, the answers were scored using 3 different marking schemes.

Scheme A

MCQs 100T/F. 'Informed negative marking but no negative marking'

The instructions for the examination were as follows

MCQs 5th MBBS Mock Examination

Date: 29th June, 2017.

Time Allowed: 1hour 15minutes

SECTION A: Numbers 1 – 20. Each question has 5 stems. Indicate whether each stem is true or false. Each correctly answered stem carries +1 mark, each wrongly answered stem carries -1 mark while each unanswered stem does not carry any mark.

SECTION B: Numbers 21 – 90. Each question has 5 stems. For each question choose the most appropriate answer; only one answer is correct. Marking more than 1 stem invalidates the answer.

In this session, the assessment was conducted as part of 5th MBBS Mock Examination. SECTION A contained 20 questions, and each has 5 stems, bringing the total number to 100 T/F questions and 70 BOF questions. For the 100 T/F questions, the candidates were informed there would be negative marking at the time of taking the examination. However, in the marking, no negative marking was implemented.

The overall scores were determined as

$$A = ([\text{no of correct answers (T/F)} + \text{no of correct (BOF)}] \times 100) / 170$$

Scheme B

MCQs 100T/F. "Informed no negative marking and no negative marking".

The candidates were given another assessment 7 days later designed to serve as part of a Continuous Assessment that would be used as a part of the 5th MBBS Main Examination. However, the same questions were administered to them.

MCQs 5th MBBS Quiz

Date: 6th July, 2017.

Time Allowed: 1hour 15minutes

SECTION A: Numbers 1 – 20. Each question has 5 stems. Indicate whether each stem is true or false. Each correctly answered stem carries +1mark, each wrongly answered stem carries no mark; similarly, each unanswered

stem does not carry any mark.

SECTION B: Numbers 21 – 90. Each question has 5 stems. For each question choose the most appropriate answer; only one answer is correct. Marking more than 1 stem invalidates the answer.

For the 100 T/F questions, the candidates were informed there would be no negative marking at the time of taking the examination. In the marking, no negative marking was implemented.

The overall scores were determined as

$$A = ([\text{no of correct answers (T/F)} + \text{no of correct (BOF)}] \times 100) / 170$$

Scheme C

MCQs 100T/F. Informed negative marking and negative marking.

This was the same examination as Session A. The candidates were informed there would be negative marking and negative marking, taken as +1 mark for each correct answer and -1/2 mark for each wrong answer, was implemented in the marking.

The overall scores were determined as

$$A = ([+CA - (WA)1/2(T/F) + \text{no of correct BOF}] \times 100) / 170$$

CA=correct answers T/F

WA=wrong answers T/F

The mean values of the scores, Positive scores and negative scores for the 3 Sessions (A, B and C) were obtained and the results compared between the groups.

Data Analyses

The data were analyzed using the SPSS version 21. Descriptive statistics were used to determine the mean, range, median, variance and percentile of the students' scores, the positive scores (correct answers) and negative scores (wrong answers) and the overall scores (mean scores). The mean values of the scores for the 3 marking schemes were compared using ANOVA. Post hoc analysis was done using Turkey/LSD. Continuous variables were compared between the groups using student t-test. Correlation statistics were used to determine the association between the marking schemes mean scores, while bivariate linear regression analysis was used to determine the strength of the mean scores for Schemes B and C to predict the mean score for Scheme A. All tests were two-tailed. P<0.05 was taken as statistically significant.

3. Results

This study evaluated the effects of negative marking in multiple choice questions in the assessment of knowledge in a set of medical students.

The students' mean scores were 46.1% for Scheme A, 57.7% for Scheme B and 39.9% for Scheme C (Table 1). The variance and percentile values are shown in Table 1. These 3 mean score values observed among the 3 marking schemes

differed significantly (Table 2). The mean score 46.1% for Scheme A was significantly higher than the mean score 39.9% for Scheme C ($p<0.001$); it was however, significantly lower than the mean score 57.7% for Scheme B, $p=0.028$.

The positive scores from correctly answered questions showed no significant difference between the 3 marking schemes, ($df=22$, $p=0.288$). Similar results were obtained for the negative scores across the 3 different schemes, ($df=20$, $p=0.862$) (Table 3).

The mean score for Scheme A showed a near perfect correlation with the mean score for Scheme C, ($r=0.97$, $p<0.001$), as well as a moderate correlation with the mean score for Scheme B (Table 3).

Multivariate linear regression analysis showed that the mean score for Scheme C predicted the mean score for Scheme A, ($p<0.001$) while the mean score for Scheme B did not predict it (Table 4).

Table 1. Descriptive analysis of the MCQs scores with 3 marking schemes

Variables	Range	Minim	Maxi	Mn/Me	SD	SE	Variance	Percentile		
								25	50	75
A Score (%)	36	29	65	46.1	9.2	1.6	84.6	39.8	46.5	54.0
A+	45	30	75	52.4				44.0	54.0	62.0
A-	30	7	37	19.4				13.8	18.5	25.0
B Score (%)	27	44	71	57.7	6.8	1.2	46.8	54.0	58.0	61.1
B+	31	52	83	68.4				63.8	69.5	73.0
B-	29	16	45	29.5				26.0	28.0	34.0
C Score (%)	35	24	59	39.9	8.7	1.2	75.8	34.5	39.5	46.1
C+	45	30	75	52.4				44.0	54.0	62.0
C-	30	7	37	19.4				13.8	18.5	25.0

A Score=Score for informed negative marking but no negative marking=correct answers. A-=wrong answers. B Score=Score for informed no negative marking and no negative marking. B+= correct answers. B-=wrong answers. C Score= Score for informed negative marking and negative marking. C+= correct answers. C-=wrong answers. Minim=minimum, Maxi=maximum, Mn/Me=mean/median, SD=standard deviation, SE=standard error of mean

Table 2. Comparison of scores between 3 marking schemes

Variables	Scheme A score: informed neg but no neg marking	Scheme B Score: informed no neg and no neg marking	Scheme C Score: informed neg and neg marking	Df	F	P value
Score % (mean \pm SD)	46.1 \pm 9.2	57.7 \pm 6.8	39.9 \pm 8.7	23	33	0.028
Positive score (Median)	54.0	69.5	54.0	22	1.398	0.288
Negative score (Median)	18.5	28.0	18.5	20	0.587	0.862

Neg=negative

Table 3. Correlation of students' mean score for Scheme A with mean scores for other marking schemes

Methods	Correlation coefficient	P value
Mean score (%) Scheme B	0.63	<0.001
Mean score (%) Scheme C	0.97	<0.001

Table 4. Multivariate linear regression analysis of students' mean score for Scheme A with mean scores for other marking schemes

Marking schemes	Beta	T	P value	95% Confidence interval
Mean score Scheme B	-0.070	-1.275	0.212	-0.244 - -0.056
Mean score Scheme C	1.022	18.708	<0.001	0.962 - 1.197

$R^2=0.951$, $df=2$, $p<0.001$.

4. Discussion

The students' mean score 46.1% in Scheme A MCQs Informed negative marking but no negative marking was significantly higher than their mean score 39.9% for Scheme

C when there was Informed negative marking and negative marking. The observed difference in the students' mean scores between the 2 marking schemes could be explained by the actual execution of negative marking in Scheme C but not in Scheme A. This agrees with earlier studies which also

showed that scores obtained by students when negative marking is adopted are usually lower than those recorded when non-negative, number right (NR) marking scheme is used [6, 13], but increase when partial knowledge is rewarded [13]. This observed difference might be explained, in part, by the fact that negative marking imposes heavier penalties for wrong answers compared to non-negative, number right schemes [6]. Furthermore, partial knowledge, assessed by a partial-credit marking scheme, in a non-negative, number right setting, tests the ability of students to discriminate between absence of knowledge and full knowledge, compared to the negative marking scheme which assesses students' ability to discriminate between absence of knowledge, misinformation and full knowledge [6, 13]. In contrast, however, some studies demonstrated no marked difference in student performance in MCQs between these two marking schemes [14, 15].

It was demonstrated in this study that the students' mean score 46.1% in Scheme A Informed negative marking but no negative marking was significantly lower than their mean score 57.7% when Scheme B Informed no negative marking and no negative marking was adopted. Specific instructions to students about marking schemes in MCQs are necessary and would influence their behavior and performance [16-18]. With clear instruction that wrong answers would attract penalties, the tendency to guess would decline [18], explaining the students' low mean score in Scheme A compared to their performance in Scheme B. Expectedly in this study Scheme A has a near perfect correlation with, and also predicted, Scheme C but did not predict Scheme B. Negative marking, as one study showed, does not address the issue of variation in risk taking behavior of students in guessing [19]. Students with equal academic ability may have different scores, as those with high risk taking tendency would be more likely to guess and to score more marks [20]. In contrast, a study has documented that guessing does not increase scores of students [16], contrary to the observation in Scheme B, in our study, in which the absence of elements of penalty perhaps, made the students mean scores to be high, obviously with instruments of guessing from the students.

A fair negative marking should employ a penalty for a wrong answer and still address the issue of guessing to be able to improve on the reliability and validity of the test method [6, 21]. In this study, -1/2 mark was adopted for each wrong answer in Scheme C, differing from those that used -1 mark [22]. The performance of a negative marking scheme which considers the elimination of guessing is dependent on the number of stems in each question; the tendency to eliminate guessing increases as the number of stems increases [5, 10, 14, 22]. A 5-stem true/false MCQ logically has 2 stems each, like ours in this study. The negative marking Scheme C in this study is fair. Outright utility of a stiffer penalty of -1 mark for a wrong answer would be of immense importance in conditions where accuracy is needed to avert dire consequences as in the medical sciences and core engineering [23, 24]. Nonetheless, Scheme C would be a fair negative marking scheme in undergraduate medicine

examinations [8].

Overall, this study has demonstrated that fair negative marking schemes adopting a penalty of -1.2 mark for a wrong answer in a structure of true/false is reliable and valid compared to a similar structure without instruction for negative marking and without implementation of negative marking. It further demonstrated a variance 84.6 and 75.8 respectively between Scheme A and Scheme C. Scheme A might be used when the scores are low and Scheme C when they are high [25].

5. Conclusions

Marking Schemes for MCQs "Informed negative marking but no implementation of negative marking" and "Informed negative marking and implementation of negative marking" were closely related, and were more reliable, valid negative marking schemes than Marking Scheme "Informed no negative marking and no implementation of negative marking" for the true/false MCQs. Negative marking should be adopted for the true/false options in MCQs for a more objective, reliable and valid assessment of knowledge in medical schools.

Limitations of the Study

The effects of +1 mark for a correct answer and -1 mark for a wrong answer were not explored in this study. The best cut-off scores were also not evaluated. These would have helped in further analyzing the effects of the marking schemes examined in this study.

REFERENCES

- [1] Mukhtiar Baig, Syeda Kauser Ali, Sobia Ali, and Nighat Huda. Evaluation of Multiple Choice and Short Essay Question items in Basic Medical Sciences. *Pak J Med Sci*. 2014 Jan-Feb; 30(1): 3–6. doi: 10.12669/pjms.301.4458.
- [2] Richard J Hift. Should essays and other "open-ended"-type questions retain a place in written summative assessment in clinical medicine? *BMC Medical Education* 2014;14:249. <https://doi.org/10.1186/s12909-014-0249-2>.
- [3] Edwina Higgins and Laura Tatham. Learning and Teaching in Action 2003, Vol 1, issue 2 Accessed on 18/09/2017 at <http://www.celt.mmu.ac.uk/ltia/issue4/higginstatham.shtml>.
- [4] Nazeem Ishrat Siddiqui, Vinayak H. Bhavsar, Arnav V. Bhavsar, and Sukhwant Bose. Contemplation on marking scheme for Type X multiple choice questions, and an illustration of a practically applicable scheme. *Indian J Pharmacol*. 2016 Mar-Apr; 48(2): 114–121. doi: 10.4103/0253-7613.178836.
- [5] Karandikar RL. On multiple choice tests and negative marking. *Curr Sci*. 2010; 99:1042–5.

- [6] Lesage E, Valcke M, Sabbe E. Scoring methods for multiple choice assessment in higher education – Is it still a matter of number right or negative marking? *Stud Educ Eval.* 2013; 39:188–93.
- [7] Richard F. Burton. Multiple choice and true/false tests: reliability measures and some implications of negative marking. *Journal Assessment & Evaluation in Higher Education.* Volume 29, 2004 - Issue 5, Pages 585-595 | Published online: 14 Sep 2010.
- [8] Ikenna Kingsley Ndu, Uchenna Ekwochi, Chidiebere DI Osuorah, Isaac Nwabueze Asinobi, Michael Osita Nwaneri, Samuel Nkachukwu Uwaezuoke, Ogechukwu Franscesca A madi, Ifeyinwa Bernadette Okeke, Josephat Maduabuchi Chinawa and Casmir James Ginikanwa Orjioke. Negative Marking and the Student Physician—A Descriptive Study of Nigerian Medical Schools. *Journal of Medical Education and Curricular Development.* 2016;3 165–170
doi:10.4137/JMECD.S40705.
- [9] Bauer, D., Holzer, M., Kopp, V. and Fischer, M.R. (2011) 'Pick-N multiple choice exams: a comparison of scoring algorithms' *Advances in Health Science Education* 16 pp.211-221.
- [10] Burton, R. F. (2001). Quantifying the effects of chance in multiple choice and true/false tests: question selection and guessing of answers. *Assessment & Evaluation in Higher Education*, 26(1), 41–50.
- [11] Burton, R. F. (2002). Misinformation, partial knowledge and guessing in true/false tests. *Medical education*, 36, 805–811.
- [12] Zimmerman, D. W. & Williams, R. H. (2003). A new look at the influence of guessing on reliability of multiple-choice tests. *Applied Psychological Measurement*, 27(5), 357 – 371.
- [13] A. Elizabeth Bond, Owen Bodger, David O. F. Skibinski, D. Hugh Jones, Colin J. Restall, Edward Dudley, and Geertje van Keulen. Negatively-Marked MCQ Assessments That Reward Partial Knowledge Do Not Introduce Gender Bias Yet Increase Student Performance and Satisfaction and Reduce Anxiety. *PLoS One.* 2013; 8(2): e55956.
doi: 10.1371/journal.pone.0055956.
- [14] Muijtjens, A. M. M., van Mameren, H., Hoogenboom, R. J. I., Evers, J. L. H. & van der Vleuten, C. P. M. (1999). The effect of a 'don't know' option on test scores: number-right and formula scoring compared. *Medical Education*, 33, 267 – 275.
- [15] Kurz, T. B. (1999). A review of scoring algorithms for multiple-choice tests. Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio, TX. Moss, E. (2001). Multiple-choice questions: their value as an assessment tool. *Current Opinion in Anesthesiology*, 14, 661–666.
- [16] Betts, L. R., Elder, T. J. -, Hartley, J. & Trueman, M. (2009). Does correction for guessing reduce students' performance on multiple-choice examinations? Yes? No? Sometimes? *Assessment & Evaluation in Higher Education*, 34(1), 1–15.
- [17] Frary, R. B. (1989). Partial-credit scoring methods for multiple-choice tests. *Applied Measurement in Education*, 2(1), 79–96.
- [18] Hammond, E. J., McIndoe, A. K., Sansome, A. J. & Spargo, P. M. (1998). Multiple-choice examinations: adopting an evidence-based approach to exam technique. *Anesthesia*, 53, 1105–1108.
- [19] Bar-Hillel, M., Budescu, D. & Attali, Y. (2005). Scoring and keying multiple choice test: a case study in irrationality. *Mind & Society*, 4, 3–12.
- [20] Choppin, B. H. (1988). Correction for guessing. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: an international handbook* (pp. 384–386). Oxford: Pergamon Press.
- [21] Espinosa, M. P. & Gardezabal, J. (2010). Optimal correction for guessing in multiple-choice tests. *Journal of Mathematical Psychology*, 54(5), 415–425.
- [22] Burton, R. F. (2004). Multiple choice and true/false tests: reliability measures and some implications of negative marking. *Assessment & Evaluation in Higher Education*, 29(5), 585–595.
- [23] Bradbard, D. A., Parker, D. F. & Stone, G. L. (2004). An alternate multiple-choice scoring procedure in a macroeconomics course. *Decision Sciences Journal of Innovative Education*, 2(1), 11–26.
- [24] Jennings, S. & Bush, M. (2006). A comparison of conventional and liberal (free-choice) multiple-choice tests. *Practical Assessment, Research & Evaluation*, 11(8).
- [25] Sanju Gajjar, Rashmi Sharma, Pradeep Kumar, and Manish Rana. Item and Test Analysis to Identify Quality Multiple Choice Questions (MCQs) from an Assessment of Medical Students of Ahmedabad, Gujarat. *Indian J Community Med.* 2014 Jan-Mar; 39(1): 17–20.
doi: 10.4103/0970-0218.126347.