

AFRICAN JOURNAL OF MEDICINE and medical sciences

VOLUME 29, NUMBER 2, JUNE 2000



EDITOR:
B. O. OSOTIMEHIN

ASSISTANT EDITOR:
A. O. UWAIFO

ISSN 1116 — 4077

A comparative study of students' performance in preclinical physiology assessed by short and long essays

EO Adewoye, ODD Oyebola and EA Bamgboye*

Dept. of Physiology and *Sub- Dept. of Epidemiology, Medical Statistics & Environmental Health, Dept. of Preventive and Social Medicine, College of Medicine, University of Ibadan, Ibadan, Nigeria.

Summary

The performance of 540 medical (MBBS) and 55 dental (BDS) students in short essay questions (SEQs) in preclinical physiology was compared with their performance in long essay questions (LEQs). The cohort was made up of 88 repeating and 452 non-repeating MBBS students and 12 repeating and 43 non-repeating BDS students. The SEQs and LEQs used for the study were those administered to the students in the paper II of the Part I MBBS/BDS examination. The results showed that all the students did significantly better in LEQs than in SEQs. When the students were sub-divided into repeaters and non-repeaters, performance in LEQs was still significantly better in all cases, except in the BDS repeaters where the difference was not significant. The study also showed that the BDS students did significantly better than the MBBS students in LEQs. Further analysis showed that the latter was due to a much better performance in LEQ 4 by the BDS students. In the SEQs, MBBS non-repeaters did better than repeaters while the situation was reversed with the BDS students. For all categories of students, on no occasion was the mean score in the LEQs or the SEQs up to 20.0, which is the 50% score and pass mark in this examination. The likely reasons for the better performance of the students in LEQs than in SEQs were discussed. The better performance of the BDS students in LEQs is believed to be due to the double exposure time of the BDS students to the topics from which LEQ 4 is drawn. From our results, it was concluded that LEQs are more useful than SEQs in assessing these students. In addition, the highly significant cross-correlation between scores in questions 1, 2, 3 and 4 suggested that the inherent problem of examiners' subjective judgement in essay marking could be minimized in both SEQs and LEQs. Finally, failure to achieve a mean score of 50% (pass mark) in both SEQs and LEQs suggested that the students are weak in essay writing. Possible reasons for this were suggested.

Keywords: Medical students, preclinical physiology, long essay, short essay

Résumé

La performance de 540 étudiants en médecine (MBBS) et de 55 étudiants en chirurgie dentaire (BDS) sur les réponses sous forme de courte dissertation (SEQs) avait été comparée à leur performance sur les réponses de questions sous forme de longue dissertation (LEQs). Ces étudiants étaient tous en cycle pré-clinique de physiologie. La cohorte avait été constituée de 88 étudiants redoublant la classe, 452 étudiants non-redoublants en Médecine, 12 étudiants

redoublant et 43 étudiants non-redoublants de BDS. Les SEQs et les LEQs utilisées pour l'évaluation avaient été les celles administrées pour la 2^e partie des examens de MBBS et BDS. Les résultats ont montré que tous les étudiants ont significativement fait mieux LEQs par rapport aux SEQs. Lorsque les étudiants étaient subdivisés en sous-groupe de redoublants et non-redoublants, la performance pour les LEQs était toujours meilleure dans tous les cas, à l'exception des étudiants BDS où la différence n'était pas significative. Cette étude avait aussi montré que les étudiants BDS avaient significativement mieux fait que ceux de MBBS en LEQs. Une analyse plus profonde avait montré que la bonne performance des étudiants BDS en LEQs était due à la 4^e question de LEQs. L'analyse des SEQs avait montré que les étudiants de MBBS non-redoublants avaient fait mieux que les redoublants lorsque la situation était inverse chez les BDS. Pour toutes les catégories d'étudiants, à aucune occasion la moyenne des scores en LEQs ou en SEQs n'avait atteint les 20.0 qui correspondent au 50% nécessaire pour réussir aux examens. Les raisons éventuelles pour la meilleure performance des étudiants en LEQs, par rapport aux SEQs, ont été discutées. La meilleure performance des étudiants BDS en LEQs pourrait être due au double temps d'exposition des étudiants au sujet de la question 4 de LEQs. À partir de nos résultats, il a été conclu que les SEQs sont moins utiles que les LEQs pour l'évaluation des étudiants MBBS et BDS. De plus, la corrélation croisée fortement significative entre les notes des questions 1, 2, 3 et 4 suggère que le problème inhérent au jugement subjectif dans l'attribution des notes pourrait être minimisé dans les SEQs et LEQs. Enfin, le fait que les étudiants n'atteignent pas la moyenne de 50% (note requise pour réussir) en SEQs et LEQs suggère que les étudiants sont faibles en dissertation. Les raisons possibles de cette faiblesse ont été suggérées. Maximum pendant les mois de Septembre à Janvier (Faison fèche). Il y avait une réduction significative. Sur le contenu d'antraquinone pendant la raison de pline. Le diagnostic microscopique de cepparance des failles dans cette petite étude est aussi décrite. Cette étude a fourni des informations vitales. Sur la meilleure des période de recotte et la distribution rationnière de l'antraquinone dans les femelles des espèce étudiées.

Introduction

The University of Ibadan Medical School, established in 1948, is the first medical school in Nigeria. The actual medical training starts in the preclinical departments where students study the basic medical science subjects in Anatomy, Biochemistry and Physiology for three semesters. At the end of this period, the students sit for the Part I MBBS/BDS examination. From 1948 to 1978,

Correspondence: Prof. D.D.O. Oyebola, Dept. of Physiology, College of Medicine, University of Ibadan, Ibadan, Nigeria.

the format of the Part I examination consisted of two long essay type question papers, each paper lasting three hours and a practical examination. The examinations in Anatomy and Physiology also include a *viva voce*. A student must pass the three subjects to proceed to the clinical stage of the course. Failure in one subject leads to a reference in that subject while failure in two or three subjects leads to repeating a year and the entire examination. A student who repeats the year and fails two or three subjects again is required to withdraw from the course.

Unfortunately, the traditional "long essay" modes of examination in medical schools have been found to suffer from several defects [1-3]. Some of these defects are that long essays test only a limited area of the subject, they consume excessive time in assessment, especially with large students' enrolment, scores obtained are highly dependent on the examiners' subjective judgement and there is poor correlation between marks awarded by different examiners for the same essays. Attempts to solve these problems led to using an increased number of questions as short essays or questions requiring only brief answers. The latter development was due largely to the reports of studies, which showed that the error in the marking of long essay questions diminishes progressively as the questions are shortened [4-6]. Also, in an attempt to solve essay-associated problems, multiple choice item examinations were introduced and they have been found to be a highly reliable method of grading all levels, from primary school to higher degree standard, including the membership of the Royal college of Physicians and Surgeons of the United Kingdom [7,8].

In keeping with the practice in many medical schools worldwide, in 1979, the format of the Part I examination at Ibadan Medical School was changed. In Anatomy and Physiology, a Paper I that was entirely multiple choice questions (MCQs) and a Paper II that was made up of short essays and long essays were introduced. In Biochemistry, short essays/questions requiring short answers and long essays were introduced.

Some studies have compared students' performance in MCQs and short essays in preclinical physiology [8,9,10]. While a few studies have also compared performance of medical students in short essay and long essay formats of examination in clinical medicine [4,6], and medical psychology [5], we found no publication in the literature, in spite of an extensive search, comparing students' performance in Preclinical physiology when assessed by the short and long essay formats. This lack of information and the need to assess the relative performance of our students in the short and long essay formats twenty years after they were introduced at the Ibadan medical School constitute the basis for the present study.

Materials and methods

The study was based on the actual performance in physiology of 595 students (540 medical and 55 dental) who took the Part I MBBS/BDS examination in August 1998. Among the medical students, 88 were repeaters and 452 were non-repeaters; while 12 and 43 BDS students were repeaters and non-repeaters respectively. The present study was based on the physiology part of paper II. Paper II consists of five questions; the first two questions are short essay questions (SEQs), questions 3 and 4 are long essay questions (LEQs), and questions 5 is

psychology. The psychology question was excluded. Thus, the study was based on questions 1 and 2 (SEQs) and 3 and 4 (LEQs) (see appendix). The time allowed for paper II was three hours.

Each question carried a maximum score of 20 marks, that is to say, each short note question carried a maximum score of 5 marks since the students were required to answer four SEQs per number. Each LEQ was marked by a Professor in the department and the same Professor marked all the candidates' scripts in a particular LEQ to ensure uniformity. The LEQ in question 4 for the BDS students was different from the LEQ for the MBBS students (see appendix). Apart from this, the entire examination was the same for both MBBS and BDS students. Other academic members of staff, of the grades of Lecturer II to Senior Lecturer marked the SEQs. Each examiner marked the same short note or short notes across board for all the candidates that attempted the short note question(s) assigned to him/her. Again, this was to ensure uniformity in marking of the short notes.

The scores of all the candidates were transferred into the computer and statistical tests of difference in performances in SEQs and LEQs were carried out between various groups as reflected under results. P values of 0.05 or less were taken as statistically significant.

Results

The results are shown in Tables 1 to 5. Table 1 shows that all the 595 students performed significantly better in LEQs than in the SEQs ($P < 0.0001$). A further breakdown of the students into repeaters and non-repeaters also showed better performance in the LEQs than SEQs for both MBBS and BDS students ($P < 0.01$). However, the difference in the scores between LEQs and SEQs among BDS repeaters did not reach the 5% level of statistical significance.

An examination of MBBS repeaters and non-repeaters (Table 1) shows that both groups still had significantly higher score in LEQs than in SEQs ($P < 0.05$). The non-repeaters had higher scores than the repeaters in both SEQs and LEQs, but the differences were not statistically significant ($P > 0.05$). Table 1 also shows that both the BDS repeating and non-repeating students did significantly better in LEQs than in SEQs. While there was a significant difference in the LEQs and SEQs scores of non-repeating BDS students ($P < 0.0001$), there was no significant difference in performance in LEQs and SEQs of repeating BDS students ($P > 0.1$). It is also worth noting in Table 1 that while MBBS non-repeaters scored higher than MBBS repeaters in both the SEQs and LEQs, the reverse was the case for BDS students; that is to say, BDS repeaters scored higher in both SEQs and LEQs than non-repeaters. The latter difference was more pronounced in the SEQs where the repeaters scored a mean of 18.46 (SD $\pm$ 3.38) against the non-repeaters' mean of 15.70 (SD $\pm$ 3.85). Therefore, the BDS repeaters did significantly better than the BDS non-repeaters in SEQs ($P < 0.05$).

Table 2 shows that when performance in SEQs were compared between MBBS and BDS students, and between all non-repeaters and all repeaters, there were no statistically significant differences in both cases ($P > 0.1$). However, when the SEQs' results are broken into MBBS non-repeaters versus MBBS repeaters, the former performed better than the latter but marginally exceeded the 5% level of statistical significance.

The comparisons of the performance of the students in long essay questions are presented in Table 3.

Table 1: Comparison of performance between short essay questions (SEQs) and long essay questions (LEQs) in physiology
Summary statistics of scores

Students characteristics	Mean	SD	N	t-value	p-value
All Students					
SEQs	16.92	4.82	595	4.587	0.0001*
LEQs	17.88	4.26			
All Repeaters					
SEQs	16.37	4.33	100	2.3899	0.0187*
LEQs	17.62	4.20			
All Non-Repeaters					
SEQs	17.03	4.92	495	3.9483	0.001*
LEQs	17.94	4.27			
All MBBS					
SEQs	16.98	4.91	540	3.5704	0.001*
LEQs	17.78	4.27			
All BDS					
SEQs	16.30	3.90	55	4.4348	0.001*
LEQs	18.95	4.08			
MBBS Repeaters					
SEQs	16.09	4.82	88	2.1781	0.0321*
LEQs	17.36	4.28			
MBBSNon Repeaters					
SEQs	17.15	4.99	452	2.9180	0.0037*
LEQs	17.86	4.26			
BDS Repeaters					
SEQs	18.46				
LEQs	19.50	3.38	12	1.4281	0.181
BDSNon Repeaters					
SEQs	15.70	3.85	43	4.2711	0.0001*
LEQs	18.79	4.32			

*Statistically significant

Table 2: Comparison of performance in short essay questions by students characteristics
Summary statistics of scores

Student characteristics	Mean	SD	N	t-value	p-value
Students(ALL)					
Repeaters	16.37	4.33	100	1.252	0.2111
Non-Repeaters	17.03	4.92	495		
Course					
MBBS	16.98	4.91	540	1.0015	0.317
BDS	16.30	3.90	55		
Students(MBBS)					
Repeaters	16.09	4.38	88	1.8816	0.0604
Non-Repeaters	17.16	4.99	452		
Students(BDS)					
Repeaters	18.46	3.38	12	2.2489	0.0287*
Non-Repeaters	15.70	3.85	43		

*Statistically Significant

Table 3: Comparison of performance in long essay questions by students' characteristics
Summary statistics of scores

Students Characteristics	Mean	SD	N	t-value	p-value
Students ALL					
Repeaters	17.62	4.21	100	0.6922	0.4891
Non-repeaters	17.94	4.27	495		
Course					
MBBS	17.78	4.27	540	1.9349	0.0535
BDS	18.95	4.08	55		
Students (MBBS)					
Repeaters	17.36	4.28	88	1.0042	0.3157
Non-repeaters	17.86	4.26	452		
Students (BDS)					
Repeaters	19.50	3.18	12	0.5284	0.5995
Non-repeaters	18.79	4.32	43		

The BDS students performed better than the MBBS students but the difference was not statistically significant ($P > 0.05$). Also, all other categorizations according to course and whether students were repeating or not did not show any significant difference in performance.

The two questions (3 and 4) that made up the long essay questions in Table 3 were examined separately in Table 4. There was no course difference in the students' performance in question 3, although MBBS students had a slightly higher mean score. However, the performance of the BDS students was significantly higher in question 4, with an average of 9.7 out of 20 compared with 8.0 out of 20 for the MBBS students ($P < 0.005$).

The simple linear correlation matrix of the scores in each question showed the correlation coefficients to be statistically significant ($P < 0.05$ Table 5). The highest correlation was between questions 2 and 4, followed by questions 1 and 2 (the two SEQs).

It is worth noting that in all categorizations in this study, on no occasion was the mean score in the LEQs or the SEQs up to 20.0, which is the 50% score and pass mark in either LEQs or SEQs in this examination. Also, some students left unanswered one or two of the SEQs.

Table 4: Comparison of performance of MBBS and BDS students in question 3 and question 4

	Question 3		Question 4	
Scores	MBBS	BDS	MBBS	BDS
Means	9.9	9.6	8.0	9.7
Standard Deviation	2.2	1.6	2.9	2.5
Sample size	540	55	540	55
t-value	1.057		4.036	
p-value	0.200		0.002*	

*Statistically Significant

Table 5: Simple correlation matrix for all students' scores in SEQs and LEQs

	SEQ(Q1)	SEQ(Q2)	LEQ(Q3)	LEQ(Q4)
SEQ (Q1)	1.00			
SEQ (Q2)	0.3083*	1.00		
LEQ (Q3)	0.2144*	0.2683*	1.00	
LEQ (Q4)	0.1850*	0.4138*	0.2524*	1.00

* $P < 0.001$

Discussion

Since this study was not contemplated till several weeks after the Part I MBBS/BDS examination was concluded, none of the present authors, the other teachers that marked or the students knew that this study would be conducted. Therefore, the results could not have been influenced by any form of bias.

There has been no previous study, as far as we know, comparing the performance of students in Physiology using SEQs and LEQs formats. There is therefore no other result to compare our findings with. The present results will however provide a baseline with which future similar studies can be compared. Unfortunately still, the two studies that compared short and long essays in internal medicine [6] and medical psychology [5] only presented the correlation coefficients between the marks of different examiners and not the actual marks scored. Apart from the fact that the subjects involved are not physiology, the non-inclusion of actual

marks in these publications makes any form of comparison difficult.

The results of the present study showed clearly that the students had a better achievement in long essays than in short essays. The reason(s) for this is not clear. Certain factors might have contributed to this. First, the students have about the same time to answer four short notes or one long essay question. The students therefore have more time to answer the long essays and present better answers than in the short essays. In addition, the imprecise definition of "how short" a short essay should be could affect students' performance in the SEQs. Some students may have difficulty deciding what to include and what to leave out while answering the SEQs unlike in the LEQs. It is worth noting that some students left blank spaces or scored zero in one or even two of the sub-numbers of the SEQs. This may be because they were ignorant or did not have enough time. These factors, operating singly or in combination may be responsible for the poorer performance in SEQs. Also, the better achievement of the students in LEQs suggests that LEQs are probably more useful in assessing these students in physiology than SEQs. The latter is in spite of the expected advantage of SEQs allowing the students to be examined over a wider scope of the curriculum than LEQs.

The findings that the repeating students did not perform well enough as the non-repeaters in both LEQs and SEQs is not surprising. It is more probable that most of the repeaters were weaker academically than their colleagues who passed the examination at the first attempt. The fact that the BDS repeaters performed better in SEQs than their counterpart non-repeaters could be fortuitous because in the same examination, the MBBS non-repeaters did significantly better than MBBS repeaters.

The better performance of the BDS students in the whole LEQs is most probably due to their better performance in LEQ 4 since there was no difference between BDS and MBBS students in LEQ 3. This better performance in LEQ 4 may be explained by the double exposure time that the BDS students have on the topics

from which the LEQ 4 is drawn. First, they are taught with the MBBS students in the regular lectures and again in a series of "oral physiology" lectures for BDS students only. The BDS students are aware that one of these "oral physiology" topics will be their LEQs 4 and they usually prepare specially for this question.

Although the SEQs and LEQs were marked by different examiners, the highly significant cross-correlation between the scores in questions 1, 2, 3 and 4 (Table 5) suggests a consistent standard in the marking of the scripts. The latter suggests that the inherent problem of examiners' subjective judgement in essay marking seem to be minimized in both SEQs and LEQs in the present study. The higher correlation coefficient between questions 2 (SEQs) and 4 (LEQs) could be due to the nature of the questions that most probably require recall of basic facts.

The poor performance of most of the students in both LEQs and SEQs suggests that the students are weak in writing essays. This may be due to inadequate knowledge of the subject and/or inability of the students to recall facts, organize them in a logical sequence, and present them as well-written essays. It is unlikely that the type of questions asked contributed to this poor performance because most of the LEQs and SEQs (see appendix) are of the factual recall type. The latter has been found to be easier to handle by students than the other types of question [11]. A recent study from our department showed that the performance of our students in Physiology MCQs was far superior to their performance in SEQs [12]. We believe therefore that many of the students who eventually passed physiology most probably compensated for their poor performance in the four questions used in this study by a better performance in the MCQs (Paper I), question 5 of paper II (psychology), the continuous assessment tests, and the *viva voce*. Scores in these other parts of the examination are not available to us, therefore, we are unable to determine their effect on the final performance of the students.

Acknowledgement

We are most grateful to all other examiners who also marked the papers

Appendix

- i. Answer all questions
- ii. Answer each question in a separate answer booklet.
- iii. Write your *Matric No.* clearly on each answer booklet.
1. Write short notes on *four* of the following:
 - (a) Cerebellum
 - (b) mass reflex
 - (c) jugular venous pressure
 - (d) gastric juice
 - (e) synaptic transmission
2. Write short notes on *four* of the following:
 - (a) blood buffers
 - (b) ejaculation
 - (c) transcellular fluid
 - (d) visual acuity
 - (e) inverse stretch reflex
3. Give an account of the control of the endocrine system by the hypothalamus.
4. (i) Define homeostasis. Write an essay on calcium homeostasis (BDS)
(ii) Give a detailed account of the neural control of breathing (MBBS).

References

1. Bull, GM., An examination of the final examination in Medicine. *Lancet* 1956, 2, 368-372.
2. Lennox B, Anderson JR. and Moorhouse, P. A comparative trial of objective papers and essay papers in pathology and bacteriology class examinations. *Lancet* 1957, 2, 396-402.
3. Nunnally, JD. In: Educational measurement and evaluation. Chapter 6, McGraw-Hill, New York, 1964.
4. Hubbard, JP and Clemans, W. V. In: Multiple Choice Examination in Medicine. Kimpton, London, 1961.
5. Mowbray, RM, and Brian M. Davies. "Short note" and "essay" examinations compared. *Br. J. Med. Educ.* 1967, 1, 356-358.
6. Ferguson A, Wright MA and McNicol GP. Appraisal of student performance: multiple-choice questions, essays, and "short notes". *Br. J. Med. Educ.*, 1970, 5, 147-151.

7. Owens SG, Robson MG, Sanderson PH, Smart GA and Stokes JF. Experience of multiple choice Question examination for Part I of the M.R.C.P. *Lancet* 1967, 2, 1034-1038.
8. Lipton A. and Huxham, G. J. Comparison of multiple choice and essay testing in pre-clinical physiology. *Br. J. Med. Educ.*, 1970, 4, 228-238.
9. Huxham GJ, Lipton A. and Hamilton D. Reasons for differential performance in multiple choice and essay tests. *Br. J. Med. Educ.*, 1975, 9, 264-272.
10. Hettiarachi ESG. A comparison of student performance in two parallel physiology tests in multiple choice and short answer forms. *Med. Educ.*, 1978, 12, 290-296.
11. Kale, OO and Bamgboye EA. Survey of short essay type MBBS examination questions in Preventive and Social Medicine at the University of Ibadan, Nigeria, (1975 – 1979). *W. Afr. J. Med.*, 1984, 3, 17-23.
12. Oyebola DDO, OE. Adewoye, Iyaniwura, JO, Alada ARA, Fasanmade A and Raji YA comparative study of students' performance in Pre-clinical physiology assessed by multiple choice and short essay questions. *Afr. J. Med. & Med. Sci.* (In press).