South African Medical Journal : Suid-Afrikaanse Tydskrif vir Geneeskunde

Cape Town, 23 April 1960 Volume 34 No. 17 Deel 34 Kaapstad, 23 April 1960

THE WRITTEN PAPER OF THE FINAL M.B., B.Ch. EXAMINATION IN MEDICINE

G. A. Elliott, M.D., F.R.C.P., *Professor of Medicine, University of the Witwatersrand, Johannesburg*

For many years I have shared the misgivings of others that the written paper in Medicine of a final qualifying medical examination affords a poor indication of the candidate's knowledge and ability to think, correlates badly with the results of the examination as a whole, including the clinical and oral sections, and does little more than separate a class into a very few bad candidates, still fewer very good candidates, and a very large group with marks bunched closely together around an indeterminate mean mark.

At the University of the Witwatersrand, written papers were abolished in the final examination in Medicine, Surgery, and Obstetrics and Gynaecology in 1952, and a statutory written examination was introduced at the end of the 5th year of study. In November 1959, written examinations were reintroduced into the final examination on account of curriculum arrangements. Certain subjects such as Gynaecology and Paediatrics were taught to a large extent in the 6th year and written papers at the end of the 5th year could not take this 6th-year teaching into account. The examination in these two subjects at the end of the 6th year was clinical and oral only, as it was in all subjects of the final examination.

In Medicine, irrespective of whether or not there is a written examination, each candidate has a clinical examination lasting 2 - 3 hours, during which he sees 3 examiners independently and examines and discusses 4 - 8 cases, some 'major' and some 'minor'. There is no exact timing system. The oral examination consists of 2 periods of 12 minutes each, during which he sees a different group of 3 examiners, 1 alone and the other 2 together. Before the written papers were abolished, the clinical and orals were regarded as carrying far more weight than the written papers, and this applies now, too, since the reintroduction of written papers. The examiners are not aware of the marks awarded for any section of the examination until a candidate has completed the whole examination.

The Department of Medicine was impressed by the advantages accruing to a final examination with no written papers, and was not anxious to reintroduce papers into the final examination, but abided by the majority decision of the Faculty of Medicine. No falling off in standards was observed by internal or external examiners during the period of abolition of written papers in the final examination.

When written papers were reintroduced to the final examination in November 1959, the Department of Medicine decided to alter the character of the written examination by introducing papers of 'objective' type additional to the papers of conventional 'subjective' or 'essay' type. The answer to the question of *purely* objective type can be only right or wrong, and marking can in fact be achieved by a punch-card system. Take, for example, the question, 'The action of insulin is to raise/lower the blood sugar. Delete which is false'. To this the answer can be only one of two—either the word 'raise' or the word 'lower' has to be deleted by the candidate. An example of the conventional essay type of question is, 'Describe the causes, clinical features and differential diagnosis of acute glomerular nephritis'. Such a question may be answered well or badly in many ways, and the subjective feelings of the examiner are prominent in the marking.

Two questions were introduced in General Medicine that were dominantly but not entirely objective. The remainder of the written questions were of the conventional essay type. There were thus 2 questions of objective type and 2 of essay type in General Medicine, 2 of essay type in Paediatrics, 1 in Dermatology and 1 in Fevers.

The two questions of objective type were as follows:

Question 3

(*a*) Name the bedside tests of assistance in the investigation of a patient with purpura.

(*b*) Define the term 'haemoglobinopathy'. Give three examples.

(*c*) List *three* drugs which can produce a haemolytic anaemia and indicate *briefly* how they exert such an effect.

(*d*) Give the rationale behind the direct Coombs test in not more than five sentences. Name *three* diseases in which the test may be positive.

(*e*) Draw short family trees illustrating the mode of inheritance of (*a*) Haemophilia, (*b*) Congenital spherocytosis.

(*f*) Give a short classification of the lymphomas.

(*g*) *Name* the treatment of choice in each of the following conditions. (No discussion is required): (i) Pernicious anaemia of pregnancy. (ii) Acute lymphatic leukaemia. (iii) Congenital spherocytosis. (iv) Chronic myeloid leukaemia. (v) Idiopathic haemolytic anaemia. (vi) Haemophilia. (vii) Christmas disease. (viii) Paroxysmal nocturnal haemoglobinuria.

(*h*) Indicate briefly the ways in which vitamin-B12 deficiency may develop.

(*i*) List *five* important clinical features that would make you suspect a diagnosis of pernicious anaemia.

(*j*) Define secondary polycythaemia. Give two important causes of this condition.

Question 4

(A) *Insulin*

(*a*) Enumerate three commonly used types of insulin.

(*b*) Record in graphic form their action on the blood sugar.

(B) *Congestive Cardiac Failure*

(*a*) Enumerate the signs of congestive cardiac failure.

(*b*) Discuss any *one* of the signs that you have enumerated.

(C) *Neurological Localization.* State sites of Lesion In:

(i) Subacute combined degeneration of the cord. (ii) Tabes dorsalis. (iii) Progressive muscular atrophy. (iv) Wernicke's encephalopathy. (v) Guillain-Barré syndrome. (vi) Trichlorethylphosphate poisoning. (vii) Lead poisoning. (viii) Diabetic neuropathy. (ix) Jacksonian epilepsy. (x) Freidreich's ataxia.

Each of these objective questions was marked independently by two examiners, neither of whom was aware of the mark awarded by the other. Examiner A set and marked question 3 and the same question was independently marked by examiner B. Examiner C set and marked question 4

and the same question was marked independently also by examiner B.

Two questions of essay type on subjects of common and general interest were selected, to be marked by two examiners independently. These questions were:

Question 1

Describe the causes, symptoms, signs and laboratory diagnosis of obstructive jaundice indicating in each group the contrast or differential diagnosis from hepatocellular jaundice.

Question 5

(*Paediatrics*)

Describe the symptoms, signs, diagnosis, laboratory investigations and treatment of acute rheumatic fever.

Question 1 was set and marked by examiner D and marked independently by examiner C. Question 5 was set and marked by examiner E and marked independently by examiner F.

For the purposes of this comparative study, the marks awarded by the examiners who set and marked the following 3 questions of essay type additional to those referred to above were selected for analysis:

Question 2: set by examiner H

Give (*a*) the essential features for the diagnosis of nephritis; (*b*) the treatment of epilepsy; (*c*) the prognosis of bronchial asthma.

Question 6 (*Paediatrics*): set by examiner E

Discuss the approach to the diagnosis of the causes of vomiting in the first year of life.

Question 7 (*Dermatology*): set by examiner G

Discuss the treatment of: (*a*) the seborrhoeic diseases of the hairy scalp; (*b*) acne vulgaris.

The marks awarded in Dermatology were selected for analysis because it was thought that Dermatology might be a subject that does not receive the same attention and interest of the candidates as subjects in General Medicine and Paediatrics.

In the total final examination in Medicine, including clinicals, orals and written papers, there were 14 failures, each with a mark below 50%; 43 passes, each with a mark of 50% - 59%; 14 exemption passes, each with a mark of 60% - 64%; and 8 second-class passes each with a mark of 65% - 74%.

For each written question under analysis, the mean mark awarded by the examiners was calculated for the candidates falling into each of these above-mentioned 4 groups in the total examination, i.e. fail, ordinary pass, exemption pass, and second-class pass. Thus, for question 3, for example, the mean mark of the 14 candidates who failed the total examination was calculated and found to be 38%, of the 43 ordinary passes 55%, of the 14 exemption passes 60%, and of the 8 second-class passes 68%. The marks of each question were analysed in this way and are tabulated in Table I and set out in graphic form in Fig. 1. From these results the following observations can be made:

TABLE I. MARKS FOR WRITTEN QUESTIONS OF SUBJECTIVE AND OBJECTIVE TYPE RELATED TO RESULT OF TOTAL EXAMINATION IN MEDICINE, WRITTEN, CLINICAL AND ORAL

Result of total examination with numbers in each pass or fail category	*Subjective questions*					*Objective questions*	
	GM 1	*GM 2*	*Derm.*	*Paed.*		*GM 3*	*GM 4*
	D	*H*	*G*	*E*	*E*	*A*	*C*
Fail, 14	52	52	37	52	52	38	43
Pass 50-59%, 43 ..	56	57	44	57	59	55	53
Exemption 60-64%, 14	59	58	47	58	60	60	64
2nd class 65-74%, 8 ..	62	59	49	60	63	66	69

Marks = mean percentage for written questions. See text for method of calculation GM = General Medicine. Capital letters identify examiners.

1. Only in the two objective written questions did the mean mark for the total examination correlate with the mean mark for individual written questions. In Fig. 1, the graph of the mean mark of the two objective questions passes through all four of the categories in which the total examination result falls, i.e. from fail to second class.

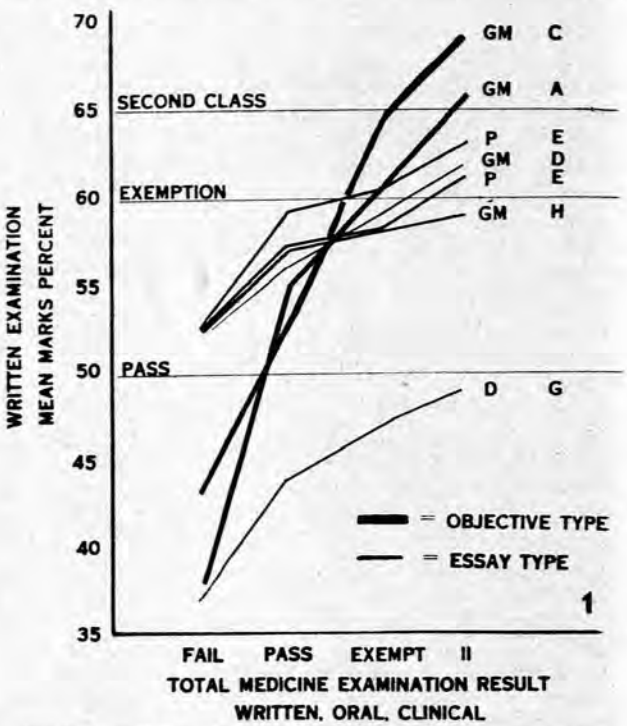

Fig. 1. Correlation of results of written and total examination in Medicine. GM = general medicine. P = paediatrics. D = dermatology. Capital letters identify examiners awarding marks.

2. In *all* the subjective questions the graph line of the mean mark ran insipidly through a small range for all categories of the total examination from failed to second-class pass.

3. In the subjective questions of *general* interest, such as jaundice, rheumatic fever, vomiting in infants, nephritis, epilepsy and bronchial asthma, this insipid graph line ran through a range of mean marks of 52% - 63%. None of the mean marks dropped to a fail level or rose to a second-class level.

4. In the subjective question in Dermatology, a subject of more limited interest, the same type of insipid graph line ran entirely below the 50% mark but about parallel to the graph line for the other subjective questions.

In other words, the correlation between marks awarded for individual written questions of objective type and the result of the total examination including clinicals and orals was good, whereas the marks awarded for the written questions of subjective type were indeterminate for fails and for the candidates with the best passes in the total examination.

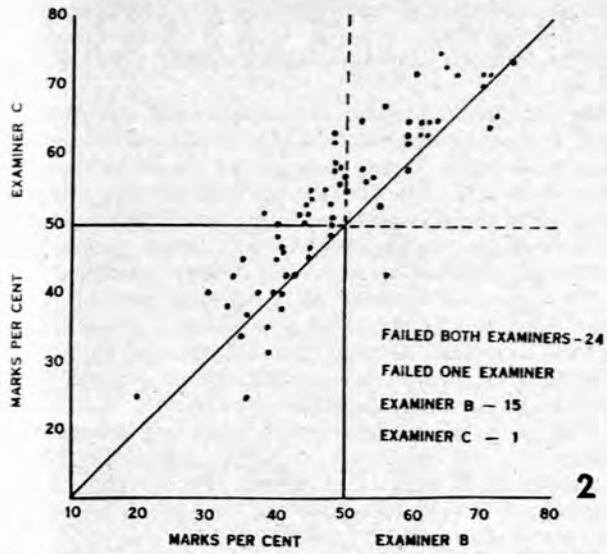

Fig. 2. Marks awarded by 2 independent examiners for written objective question (question 4).

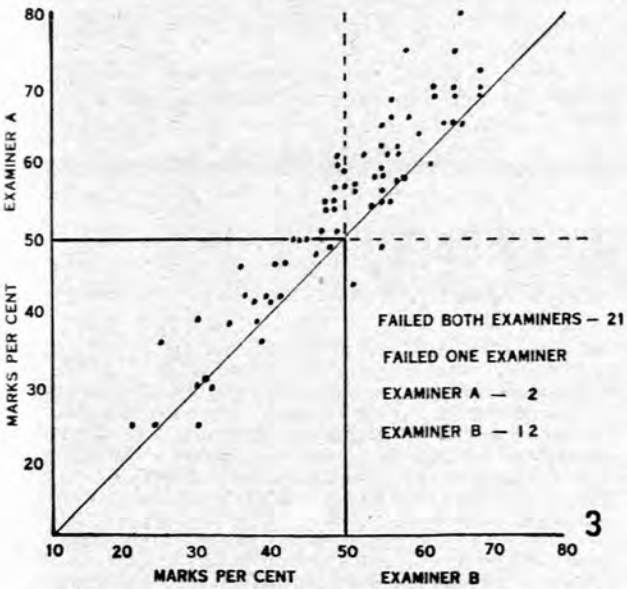

Fig. 3 Marks awarded by 2 independent examiners for written objective question (question 3).

In Figs. 2 - 5 are set out graphically the individual marks awarded by 2 independent examiners marking each of 4 questions, 2 of them objective and 2 subjective. From these diagrams the following observations can be made:

1. In the two objective questions (Figs. 2 and 3) the marks of both examiners covered a wide range from fail to second-class pass with a good scatter and good correlation in spite of individual differences in standards as shown particularly by the numbers failed by both examiners and by a single examiner.

2. In the two subjective questions (Figs. 4 and 5) the marks of both examiners were bunched together in the pass area over a narrow range and with poor scatter. Very few candidates were failed by one or both examiners.

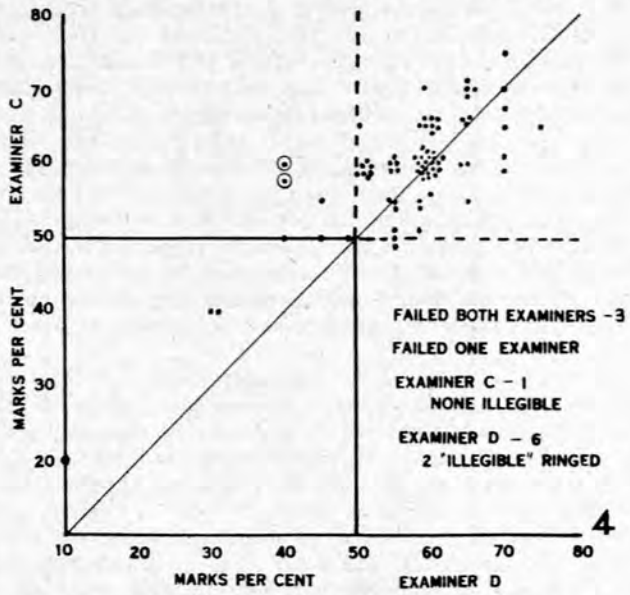

Fig. 4. Marks awarded by 2 independent examiners for written subjective question (question 1).

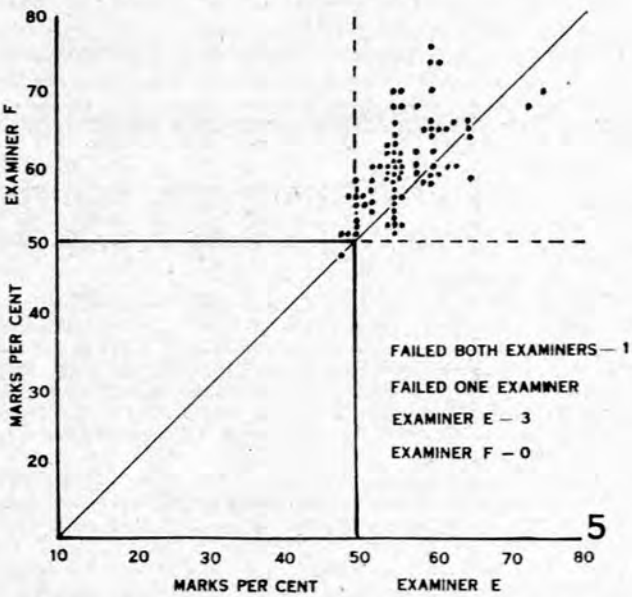

Fig. 5. Marks awarded by 2 independent examiners for written subjective question (question 5).

3. In the objective questions examiner B, who was the second examiner in both, failed considerably more candidates than examiners A and C respectively, who had set and marked the questions and whose marks appear in Table I. Clearly, the factor of the personality of the examiner enters into these questions, although not interfering with the correlation of marks between the two examiners and not interfering with the correlation with the result of the total examination. It was found in discussion that examiner B had made high demands of accuracy. For example, he marked down candidates who failed to record correctly the exact times of fall and rise of blood sugar after injection

of insulins, whereas examiner A gave more credit for graphs of blood-sugar curves that gave a general indication that the candidate knew the comparative differences in action of the three insulins even though the recording of individual times of fall and rise of blood sugars were not accurate.

4. Examiner D gave poor marks on account of 'illegibility' to two candidates, whereas examiner C, marking independently, gave good marks to these same two candidates and did not regard their handwriting as illegible. The surname initials of these two candidates were V and W, suggesting the possibility that examiner D was fatigued as he approached the last papers, and found illegible handwriting that he might well have deciphered had he been in a less fatigued state.

DISCUSSION AND CONCLUSIONS

In spite of the fact that the objective questions set in this examination were not purely objective, in that they contained discussion and essay-type elements, the marks awarded for these questions were far more discriminative than the marks for the subjective questions, and correspond well with the final total result of the examination, including results of the clinical and orals. The marks awarded for the questions of subjective type bore no relationship whatsoever to the results of the total examination.

Correlation of marks awarded by independent examiners was good for the objective questions in spite of the severer marking standards of one of the examiners. This difference in marks did not interfere with the good correlation with the final total result of the examination. Objective questions lend themselves to standardization of answers. In the recent trial no attempt was made beforehand to set the standards of marking of the objective types of question, although this could legitimately have been done and will be done in future.

In the subjective questions, not only did the marks not bear any relation to the final total examination result, but the marks given by two independent examiners bore little relation to each other owing to their all being bunched together in the pass range with very few failures. It is impossible to standardize answers for subjective questions.

The two types of question test different qualities:

The objective question demands accuracy of factual knowledge and at the same time affords the candidate a chance to express himself concisely and clearly either in words, which tests literacy, or by diagram. Without factual knowledge the candidate is bound to fail.

The essay-type question tests literacy and the power of expression, but ignorance of factual knowledge can be concealed by a well-written paper. Furthermore, a great amount of factual knowledge can be presented badly and the candidate be awarded a poor mark.

There would appear to be a place for each type of question in the final medical examination, but the one that bears correlation with the final total examination result is the objective question. The question of essay type is indeterminative.

It might, however, be justifiable to ask, if the objective written question correlates well with the final total result of the examination and the essay-type question correlates not at all, why have written questions?