

Correspondence

Multiple choice question examinations

Sir,

I was interested to read in January's *Journal* of the accuracy of the multiple choice question method of examination. As a candidate who sat the November examination I would like to say that I have difficulty in regarding it with this degree of enthusiasm. It is open to the following objections.

1. It is difficult for the candidate to acquire information about the multiple choice questions, because the examination authorities are fearful of their question papers being "compromised".

2. The title is misleading. There is no choice of question at all. The candidate is confronted with a long series of questions, each of which is compulsory. There is a choice of 5 possible answers, most of them distasteful, and with one of these he had better agree, if he wishes to score any marks at that particular examination. This applies whether he considers the correct answer to be represented in the possibilities or not.

3. The candidate is denied one of the fundamental freedoms of man. He is denied freedom to express himself and is forced into the humiliating position of having to appear to agree with that in which he may not believe. How is it possible to judge fairly a man who is not allowed to write or say what is in his mind?

4. Candidates are instructed that if they do not know the answer to a question, they should leave it undone and proceed to another question. Above all, they should not guess. All this is foreign to our training and our work situation. We cannot turn our backs on our patients just because we do not know right away what is wrong with them. We have been trained to persist to get the answer. This may come in one of the following ways.

- (a) The patient tells us what is wrong.
- (b) The doctor can see what is wrong.
- (c) The doctor gets the answer by a process of logical deduction.
- (d) If (c) fails, he falls back on his intuition, i.e. inspired guesswork, which is not allowed in the examination.
- (e) He calls for specialist help.

Whatever happens, he does not throw in the

towel just because the answer is not immediately apparent. But this is what he is asked to do in the examination. Therefore the examination is not representative of operational conditions.

5. Many of the questions were superficial and would encourage candidates to go in for a superficial type of preparation, e.g. the knowledge of the most common indication for caesarean section is not really so terribly important, but it is vital that the practitioner be aware of all the indications for caesarean section and be capable of recognizing them.

6. Some questions were so technical that they would have required the services of the appropriate consultant to answer them.

7. Other questions were awkwardly phrased and left a candidate with a full and accurate knowledge of the subject, in a quandary as to which answer to select.

For example the question on "Administration of oxygen to a new-born baby should not exceed 10 minutes, 1 hour, 3 hours, 6 hours, 12 hours." No mention was made of the mode of administration, e.g. by mask, tent or hyperbaric incubator. Professor Hutchison, of Glasgow, in his textbook, suggests 15 minutes as an average in the hyperbaric incubator and 30 minutes as a maximum. Unfortunately, the times mentioned do not coincide with those specified in the question. But then, should the results of research be tailored to the examination question or should the examination question be tailored to the results of research?

Other workers have suggested longer periods of oxygen administration but what is of considerable interest is that a few days after the written paper took place the Scottish Home and Health Department issued a report of a subcommittee of the Standing Medical Advisory Committee, entitled "Uses and Dangers of Oxygen Therapy" published by H.M.S.O. price 9/6. This report, after considering the evidence in detail, summarizes the matter by stating in paragraph 119, page 39 "Oxygen should not be given for any period longer than is necessary."

Speaking personally, I still do not know which of the alternatives was considered to be

the right answer to the above question. I do not even see how, from the knowledge available to a candidate, he could have arrived at a definitely correct decision in the terms of the alternatives put forward. He may have justifiably reckoned that 15-30 minutes did not exceed 1 hour. He may equally justifiably have decided that 15 minutes was much nearer to 10 minutes than 30 minutes is to 1 hour.

The question did not ask "Why" so much care should be taken in oxygen administration to the new-born. It was therefore possible for a candidate who was ignorant of the possibility of retrolental fibroplasia to score full marks, while another candidate, fully conversant with the danger, was penalized unjustly.

8. No-one who has spent months assiduously preparing for an examination likes to be marked by a computer. The marking of examination papers demands the highest qualities of human judgement.

9. In the nature of things at present many of the candidates for the MRCGP are not in the first flush of youth. Many are over 40 years of age and some are nearer 50. It is not reasonable to subject such people to the squiggles of computer cards for the first time in their lives in the middle of an important and, to them, expensive examination. In my own case I was so dazed by the computer cards and the speed of decision-making that I had to stop for a rest for a few minutes about the 150th question. Fatigue, visual and mental, is a problem, particularly for older candidates. They may know the correct answer but mark the wrong space on the card.

10. The attraction of the multiple choice paper is that it cuts out a lot of time consuming writing and forces the candidate to make decisions, at the rate of 200 in 120 minutes. This certainly shows up a candidate's deficiencies but does not give him adequate time to exhibit such skill as he does possess. It seems to me that a reasonable compromise would be to have a written paper consisting of say 50 short questions with the candidates being made to understand that only 2-3 lines of writing was necessary to answer each question. Verbosity could be penalized. The question on oxygen administration could then be rephrased as follows.

"(a) State briefly what precaution should be

taken in the administration of oxygen to the new born.

"(b) Why?"

The uninformed candidate would be trapped by this form of question much more effectively than by the multiple choice question. The knowledgeable candidate is allowed room to deploy his knowledge and is allowed freedom to express himself.

11. There is much concern about compromising multiple choice questions. This is not something peculiar to multiple choice question papers. Any examination paper is compromised by publication. In fact, any question is compromised the very moment it is uttered. There is therefore a tremendous challenge to examiners to find new questions and perhaps to dress up old ones. This is the way the matter should be approached rather than by suppression of publication.

12. It is stated that "A test is said to be reliable if it gives consistently similar results when taken by different groups of candidates." But surely it may be equally argued that a test which is shown to be inaccurate with one group of candidates will perpetuate its inaccuracies when repeated with other similar groups of candidates.

13. The alternatives offered in certain multiple choice questions may be compounded in such a way as to come near to violation of Section 14 of The Medical Act 1956. This section states "If it appears to the General (Medical) Council that any university or other body . . . has attempted to impose upon any candidate offering himself for examination an obligation to adopt, or to refrain from adopting, the practice of any particular theory of medicine or surgery as a test or condition of admitting him to examination or of granting a certificate, the General (Medical) Council may make representations to the Privy Council . . . (who) may direct the body to desist from such attempts. . . ." I do not suggest, for one moment, that the examiners, in this particular instance, purposely set out to inflict unacceptable views on unsuspecting candidates. However, I feel that the multiple choice question method of examination is such as eventually to compromise the sanctity of Section 14.

I feel that no educational body should subject candidates to examination questions which cannot face the light of day. Even I

who sat the examination have only recollections to support my argument. Such controversial methods of examination should be open to the closest scrutiny by all concerned, including the general public, who are ultimately affected by them. I am perturbed that an institution, with the high ideals of the Royal College of General Practitioners, should sully its hands, even temporarily, with an instrument of examination, which by necessity carries within it, the seeds of its own eventual disrepute. The mere fact that other Colleges have embarked upon this clandestine method of examination is no great recommendation to us to emulate them in their monumental folly and disregard of human rights.

In conclusion, let me say that I attack only the method of this particular part of the examination. I have only respect for the examiners, who throughout the proceedings, exhibited the highest possible standards of courtesy and must have laboured unceasingly to compensate for the shortcomings of this particular part of the system.

Glasgow

DUNCAN McNICOL

We have shown Dr McNicol's letter to the Chairman of the Board of Censors who has made this comment:

Dr McNicol raises several points, some of which, in all fairness, deserve an answer, albeit in more moderate terms than those in which they are made. Yet the greater significance lies in what he did not say rather than in what he said. There is demonstrated an unawareness of the purpose and function of the examination as a whole, which represents a lack of communication between the Board of Censors and candidates. For this, as chairman of the board, I must accept responsibility and I propose to attempt some clarification.

Our examination seeks to assess candidates in the areas of the knowledge, or better—factual recall, skills and attitudes appropriate to general practice. The different portions of the examination are each calculated to assess certain aspects only. It is also implicit that the several portions should be fair, relevant, discriminating and reliable tests. The multiple choice question paper, as we use it, seeks to measure factual recall in a variety of areas. These are set out in the report of the Conference of Examinations^{3, 4}. It thus forms only one of the four portions of the examination. While

in his ultimate sentence, Dr McNicol mentions that there are other parts of the examination, his readers could be forgiven the assumption, until they reached this sentence, that the whole examination was under attack.

The difficulties of the MCQ lie mainly in construction and validation. The techniques of construction are well described by various authors and two works in particular^{1, 2} could be consulted for details. We rely on the 'one-out-of-five' type of answer because:

1. Experience has shown it to be a simple and reliable method.

2. The candidate is spared the extra stress of interpreting the varying, and sometimes difficult, instructions which must precede each change of type of question.

It is the task of the examiners to create a 'bank' of questions, each one of which has been validated. The process of validation is as follows:

First, and perhaps the most difficult, the questions are formulated within the rules for construction referred to above. The most recent and authoritative sources are used.

Secondly, the questions are reviewed by a small group of the examiners, and any questions producing immediate disagreement are modified until agreed, or rejected forthwith.

Thirdly, the paper is 'sat' by a group of volunteers, and the papers are machine marked. This enables the examiners to determine the level of discrimination of the examination.

The detailed analysis provided of each question enables yet more to be rejected or modified.

Lastly, a minimum pass mark for this portion of the examination is set.

In the case of the last MCQ paper these steps had been taken and the original paper of 220 questions was sat by 60 volunteers—all teachers in general practice. From this paper, as a result of the question analysis, 24 questions were rejected and four new ones added, bringing the paper up to 200 questions, for which the time of two hours had been demonstrated by the volunteers to be adequate.

After the college examination to which Dr McNicol refers, some 60 questions, of which only a few were shown to have been undesirable by this further analysis, have been removed and the paper for the May examination will be of 220 questions, chosen to demand, as near as is possible, an equal level of answer.

Thus, in the MCQ papers the great labour is in creating the questions. It is at this time that the highest qualities of human judgment are