

EXAMINATIONS

Douglas M. McIntosh

Principal, Moray House College of Education, Edinburgh, Scotland.

Examinations have a determining influence on careers, and accordingly they arouse emotional attitudes: they have been referred to both as "the central nervous system of education" and as "a millstone round the necks of the schools." Such reactions arise partly because examination results are used without sufficient thought to their validity and partly because many examinations are not constructed with clearly defined objectives. Despite the early researches which revealed the extent of the unreliability of examination marks, only recently have steps been taken to produce the improvements which are so necessary in any system of large-scale examining. For example, the pass/fail examination is a weak measuring instrument because all measurements of human ability are approximate and to use an examination to distinguish between a pupil who scores 50 and is awarded a "pass" and another who scores 49 and is awarded a "fail" assumes a degree of accuracy which is beyond the resources of a written examination.

A very large subjective element enters all aspects of examinations. The selection of questions involves the sampling of sections of a syllabus on which the examinee is tested: the sampling depends entirely on the views of the examiner. Pupils have sat examinations in which they have "spotted" certain questions and, accordingly, found the examination easy, whereas others have had the experience in which much of their hard work has gone for naught because the sample has been drawn from areas of the syllabus on which they had not concentrated their attention. A subjective element is also involved in the allocation of marks. For example, a group of teachers or students when given no guidance will award a wide variety of marks to an arithmetic question in which the pupil takes down one figure wrongly but works the sum out correctly: the different marks indicate different value judgements.

Examinations have a wide range of purposes. The following are the most common:

- (a) attainment : standards and objectives must be clearly defined and the examinees should be given some indication of the standards which they are expected to attain.
- (b) diagnosis : such tests or examinations attempt to find out what the pupil does not know. For example, a diagnostic test in arithmetic may show that a pupil has a specific weakness: many young children have difficulty with zeros and when such weaknesses are revealed by a specially constructed test they can be corrected.
- (c) prediction : most of the examinations at the end of secondary school are used to predict the success with which a pupil will engage in a course of higher education. Many such examinations have not been specially constructed for this

objective, and it is merely assumed that they have this value although investigations have shown that the relationship between success in higher education and secondary school examination results is not high.

- (d) motivation : the knowledge that a specific standard has to be achieved by a certain time gives purpose to teaching and to learning. Teachers interested in their pupils who are attempting to gain admission to university will strive hard to ensure that the pupils gain the necessary entrance qualifications, and the pupils for their part have a clearly defined goal which channels their energies.
- (e) selection : selecting a number of pupils for a particular course of study, where the purpose of the examination is to ensure that pupils who are chosen have the ability and the knowledge which will enable them to undertake a particular further course of study. When a restricted number of candidates have to be selected from a group, it is assumed that the examination can make the fine distinction between the candidates at the border line.
- (f) a teaching instrument : too often, once the examination marks are issued, no further work is done on what has been written. On the other hand, examinations or tests of some nature should be a regular feature of classroom work. These need not be formal examinations but tests to check whether the learning experiences created by the teacher have been effective. In all teaching situations, the teachers should have clearly defined objectives. Following the exposure of the taught to a learning experience the teacher must attempt to assess the success with which the objectives have been achieved. The assessment will have two purposes: first, to ensure that the learning experience is effective, and secondly, to gain some insight into the ability and attainments of the taught.

One of the weaknesses of examinations is that they tend to be used for a wide variety of purposes. Examinations at the end of secondary school courses are used for entrance qualifications to a large variety of professions, which select these examinations merely because they are available. It has often been suggested that different professions should set their own examinations, but this would place the schools in an impossible situation whereby they would require to prepare students for a very wide variety of examinations. National external examinations such as the GCE in England and the SCE in Scotland are used for entrance qualifications for many courses in higher education and while they were originally intended for university entrance they now are used by a multiplicity of institutions providing higher education: it is doubtful whether they can have a high predictive value for all institutions.

An examination should test the objectives of a specified course of study, and these objectives should be clearly defined. In the case of schools, teachers should create the learning experiences which enable pupils to attain the stated objectives and the examination should provide a measure of the effectiveness of the learning experience. The knowledge of facts is essential in any course, but too many examinations rely heavily on memorisation. In an examination on statistics or mathematics pupils are allowed to use mathematical tables, but there seems no good reason why they

should not also be allowed any books which they use in the classroom. The memorisation of a formula is of little importance in life outside the school: it is the ability to use the formula which is of real significance.

Objectives have been classified by Bloom in his widely known work "Taxonomy of Educational Objectives"* and it is possible to analyse examinations showing the percentage of each category which have been tested by the examinations. The categories are:

- (a) information
- (b) understanding
- (c) application
- (d) analysis
- (e) synthesis
- (f) evaluation

These are in what is known as a hierarchical order: each category cannot be carried out without skill in the previous categories, for example, understanding is not possible without information and, similarly, application cannot be carried out without information and, similarly, application cannot be carried out without information and understanding.

Too great reliance is often placed on examination marks: as has been indicated the pass/fail examination provides an example of an assumption of the accuracy which it is assumed examinations possess. Much research has been conducted to show the unreliability of examination marks. Pupils sitting similar examinations at intervals often show a degree of unreliability which is difficult to understand. The lack of consistency in examination marks may be due to three factors:

- (a) the unreliability of the marking: research has shown that the same examiner can award different marks to the same pupils after an interval of time. In one investigation 14 examiners were asked to re-mark 15 history scripts some 12 to 19 months after the first time, having kept no record of their previous marks. The examiners awarded not only numerical marks but also the verdict of failure or pass or credit. It was found that in 92 cases out of 210 the individual examiners gave a different verdict on the second occasion from the verdict awarded on the first occasion.
- (b) the unreliability of the examination: questions in any examination can deal with only a sample of the syllabus which has been studied. If the sample is not representative of the syllabus as a whole the examinees may not do themselves justice. When a teacher can make an accurate forecast of the type of question likely to appear in an examination, his pupils will probably gain marks unrepresentative of their ability. To give consistent results an examination should be of the same order of difficulty to all examinees of comparable ability.

*

Bloom, B.S. (ed): Taxonomy of Educational Objectives, Handbook I: Cognitive Domain, David McKay & Co., Inc., New York, 1956.

- (c) the unreliability of the examinee: few individuals can consistently reach the same standard in any activity: all have their good and bad days. Some are more consistent than others and it is for this reason that a decision on a pupil's future should not be based on the result of a single examination. Ill health, or nervous tension, may also cause an examinee to do badly in an examination.

One technique which has been successfully developed to overcome the unreliability of examinations is the objective test. "The Battle of Bannockburn was fought in - 1413 1143 1314 1134 - underline the correct answer." Questions of this type have only one correct answer and thus the mark is likely to be the same no matter who corrects the paper. A large number of such questions make up a test and therefore a much wider sample of the syllabus can be studied. Such tests have some obvious disadvantages in that they cannot test a pupil's ability to select data and arrange it in an orderly and logical order. The construction of such tests also is a much more laborious and time consuming operation than an essay type of examination.

The most important factor in determining success or failure in examinations is motivation. Where pupils have a strong desire to succeed they are more likely to do well: where they have little interest in the result the standard achieved will be low. Care has to be taken, however, to make sure that there is no over-anxiety on the part of the examinees, otherwise they become upset and do themselves less than justice. One of the causes of uneven performance by examinees is the undue proportion of an examination dependent on memorisation: too many examinations cannot be answered without remembering information. For example, the following question is typical of some examinations - "How many balls of $1\frac{3}{4}$ " radius can be made from 539 cubic inches of metal?" Examinees may be capable of answering this question, but cannot remember the formula for the volume of a sphere. Outside the examination room, anyone who had forgotten such a formula would simply turn to a book for the answer. The present explosion of knowledge makes memorisation more and more unprofitable since much information becomes rapidly out of date. To find out how to use facts is much more important than remembering them, and examinations should reflect this change in emphasis.

It cannot be too strongly emphasised that all assessment of human ability is approximate, hence the more reliable and valid information about an individual which can be ingathered, the more any assessment is likely to give a true measure of the individual's ability or attainment. The result of a single examination should never be used as a means of assessment. Three components ought to be used:

- (a) teachers' assessments : the pupil's record over a number of years will give a rating which should be fairly reliable. Teachers standards, however, vary according to the teacher's experience, hence some form of external measure is necessary to scale the assessments on a uniform standard.
- (b) external tests of objective type : these give a reliable measure of a pupil's ability in a clearly defined syllabus.
- (c) essay type examinations : these measure other aspects of achievement and ability than those generally measured by the objective test.

Examinations are not the same thing as the day of judgement: and they are certainly not the Alpha and Omega of education, nevertheless some form of assessment is essential if teaching is to be effective and the potential ability of individuals is to be developed: the ruler is unlikely to be discarded because it cannot measure to 1/1000th of an inch.