

A STUDY OF OPTIONAL QUESTIONS IN EXAMINATIONS

Lindsay D. Mackay

Faculty of Education, Monash University, Melbourne, Australia

The case for inclusion of optional questions in an examination is frequently argued on the grounds that a particular examination is not designed to measure whether a student possesses knowledge or facts, but whether he has developed particular abilities or skills which can be assessed independently of the particular question or questions answered. If the assumption that these skills and abilities can be assessed independently of the particular question answered is accepted, the inclusion of optional questions in a public examination allows the teacher greater freedom to develop these skills and abilities in any of a number of sections of the broad subject area, and it allows greater freedom for the individual student to pursue his interests through independent study.

From a measurement point of view there are considerable difficulties in the use of optional questions in an examination. Many people would argue that the use of optional questions in an extended answer examination adds one further source of variability to the subjectivity and inaccuracy that already exist in extended answer examinations. If a student is permitted to choose which questions he wishes to answer, the basis for comparability of scores is considerably weakened, because different students will answer samples of questions which are not comparable in content, abilities or objectives. As a result, the content validity of the examination differs considerably for different students. If a choice of questions is available, the questions answered by a particular student are likely to provide a limited, and perhaps distorted, sample of that student's achievement in the course, because he will tend to choose those questions he is best prepared to answer, and because the availability of options gives him greater opportunity to use materials prepared by others to produce an answer which does not reflect his ability. Rather than being fairer to all students, as some advocates of optional questions in an examination would argue, the opportunity to choose among optional questions may help the poorly-prepared student considerably more than it helps the well-prepared student.

This report is a brief summary of some of the results of a study of the effects of optional questions in examinations. As part of this study, the marks awarded to optional questions in a number of examinations, ranging from the sciences to the humanities, have been analysed in an attempt to determine -

- (i) the extent to which optional questions differ in difficulty,
- (ii) the extent to which optional questions and different marker interact,
- (iii) the extent to which students of different ability select optional questions of different difficulty, and
- (iv) the extent to which optional questions differ in their marker reliability.

(i) Differences in difficulty of optional questions

One problem in estimating the difficulty of optional questions is that the groups of students attempting each optional question differ in ability. In an examination which consists of a compulsory section in addition to optional questions, it is possible to use a student's score on the compulsory section as an index of his ability in the subject being examined. It is then possible to calculate the average score on each optional question for candidates in each range of scores on the compulsory section (i.e. in each ability range), and to use these average scores to estimate the average score that would have been obtained on each optional question if all students sitting for the examination had attempted each optional question. If the examination does not contain a compulsory section, the average score obtained by the students on all other questions attempted on the examination can be used in a similar way, as an index of the student's ability in the subject of the examination.

An indication of the differences in difficulty of combinations of optional questions available to students can be gauged from the results in Table 1. In this table, the differences between the estimated average scores on the least difficult combination of questions available to students are given for six examinations.

Table 1: Differences in average scores between least difficult and most difficult combination of questions available to students in six examinations containing optional Questions.

Examination	No. of students used in the analysis	Choice available	Difference in estimated average mark between least and most difficult question combinations (expressed as a percentage of possible marks on optional questions)
1. Grade 11 Physics	4360	3 out of 5	9.7%
2. Grade 12 Physics	4450	5 out of 7	4.7%
3. A Grade 12 Humanity	2230	5 out of 14	3.5%
4. A Grade 12 Humanity	2150	5 out of 13	3.3%
5. Grade 12 History	7410	3 out of 13	6.5%
6. Under-graduate Physics	190	5 out of 11	10.2%

The entries in the table can be regarded as estimates of the differences in average marks that would have been obtained had the same candidate attempted two different combinations of questions, or had two candidates of "equal ability" attempted different combinations of questions. It is apparent that considerable differences in the results on an examination containing optional questions can arise from differences in the difficulty of the options offered, and that these can produce a difference in average marks obtained by candidates of equal ability of up to 10% of possible marks.

(ii) Interaction of examiner and optional questions

It has long been recognized that different markers mark essay questions to different standards. In the analyses of examination papers in this study it is clear that markers do not mark all questions to a consistently hard or easy standard, and there is evidence of interaction between marker and optional question which results in a marker marking different questions to different standards. An idea of the magnitude of the effect of this can be gauged from the results in Table 2, in which are given the maximum observed differences in the average marks that would be obtained by candidates of "equal ability" who attempted different question combinations, and whose scripts were marked by different markers. Results in the table are based on analyses of Examinations 3, 4 and 5 in Table 1. Examination 3 was marked by 15 markers, Examination 4 by 14 markers and Examination 5 by 31 markers. The system of allocating scripts to markers was such that each marker marked every possible question.

Table 2: Maximum observed difference between average scores obtained by students of "equal ability" as a result of a number of sources of variation.

Source of variation	Maximum observed difference between average scores obtained by students of "equal ability" as a result of this source of variation.		
	Examination 3	Examination 4	Examination 5
	%	%	%
Different questions (irrespective of marker)	3.5	3.3	6.5
Different markers (irrespective of question)	6.8	5.1	9.5
Same questions, marked by different markers	7.8	10.0	23.1
Same markers, marking different questions	10.8	9.6	18.8
Different questions, different markers	13.0	14.1	29.9

It is apparent that there is considerable interaction of question and marker and that two students of "equal ability" would be expected to obtain average marks which differed by up to 30% of possible marks, depending on the optional questions they selected and the particular marker to whom their scripts were assigned. The entries in Table 2 are the observed maximum differences in average scores between students of equal ability due to various sources of systematic variation in the marks awarded to students. Some students would experience differences considerably greater than these average values, and other would experience differences considerably less than the average values.

If a statistical correction had been applied to the marks awarded to different questions by different markers, so that students of "equal ability" would be expected to obtain the same average mark on each question irrespective of the marker who marked the question, then a considerable percentage of the students in the above examinations would have received a different final grading.

(iii) Ability of students attempting different question combinations

In an examination which consists of a compulsory section and optional questions, it is possible to use the compulsory section mark as an index of the ability of students who attempt different combinations of optional questions. As described earlier, a measure of the difficulty of the question can be obtained by estimating the average mark that would have been obtained had all students attempted the question. There is evidence that the most difficult question combinations were chosen by a group of students of significantly lower average score on the compulsory section than the group of students who chose the least difficult question combinations. Some of the results obtained are summarized in Table 3 for Examinations 1 and 6 in Table 1.

Table 3: Differences between average scores on the compulsory sections of two examinations for students attempting the least difficult and most difficult combinations of optional questions.

Examination	Difference in estimated average score of least and most difficult question combinations (as percentage of total score on all optional questions).	Average compulsory section score of students attempting <u>least</u> difficult questions minus average compulsory section score of students attempting <u>most</u> difficult questions (expressed as a percentage of possible marks).
1	%	%
6	9.7	+21.5
	10.2	+ 8.6

(iv) Differences in marker reliability of optional questions

Each of the scripts completed by the 7410 students who sat for a Grade 12 History examination (Examination 5 in Table 1) were independently marked by two of the 31 markers who marked scripts for this examination. The consistency with which the same answer was independently marked by pairs of markers for each of the 13 optional questions on this examination is expressed in three ways in Table 4. Firstly, for each question, the correlation between the marks awarded by two markers to the same answer is given. The mark/remark correlation values in the Table vary from .53 to .67 for the 13 questions on this examination. Secondly, the percentage of variance in the marks awarded to an answer by one marker which is common to the mark awarded by the second marker is given. This percentage represents the percentage of common variance in the marks awarded to the same answer by two markers. The percentage of the variance of scores which is unreliable variance or error variance can be obtained by subtracting the percentage of common variance from 100%. In the Table, the percentage of variance which is common to the two marks awarded to answers on each question varies from 29% to 46%; that is, the percentage of variance of marks which is error variance varies between 71% and 54%. The third measure of consistency in the Table is the average size of the difference of marks awarded to the same answer by two markers. The values obtained range from an average difference of 9.9% of possible marks to an average difference of 12.5% of possible marks.

Table 4: Correlation between marks awarded independently by 2 markers to answers on each of 13 questions on a Grade 12 History Examination.

Question	Correlation between marks awarded independently by 2 markers.	Percentage of variance common to the marks awarded by 2 markers.	Average difference between marks awarded to the same answer by 2 markers (as a percentage of possible marks.)
		%	%
1	.56	31	11.6
2	.57	33	12.5
3	.53	29	10.5
4	.60	36	10.5
5	.67	45	12.3
6	.55	31	10.0
7	.67	46	10.9
8	.65	43	10.5
9	.65	42	11.4
10	.63	40	9.9
11	.65	43	11.3
12	.65	43	10.4
13	.66	44	12.4

Summary

This report briefly summarises some results of a study of the effects of optional questions in six examinations. The results indicate that

- (i) Differences in difficulty of different combinations of optional questions available to students can result in average differences of between 3% and 10% of possible marks between students of equal ability who select different question combinations.
- (ii) Different markers mark different optional questions to different standards, and, as a result, students of equal ability who answer different questions and whose answers are marked by different markers could be expected to obtain average marks which differ by up to 30% of possible marks.
- (iii) The group of students who select the most difficult combinations of optional questions on an examination are of significantly lower ability than students who answer the least difficult question combinations.
- (iv) There are considerable differences in the reliability with which different optional questions are marked.

The report has indicated a number of difficulties associated with the use of optional questions in examinations. Further work is currently in progress to investigate other effects of optional questions in examinations.