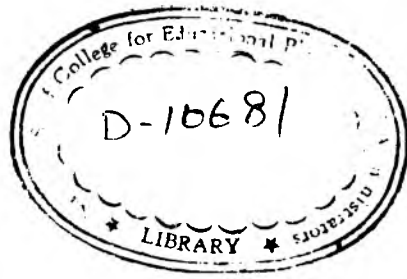


REPORT
ON
EXAMINATION REFORM

UNIVERSITY GRANTS COMMISSION
NEW DELHI



C o n t e n t s

	Preface	...	v
1	Introduction	...	1
2	Appointment of the Committee	...	4
3	Statement of the Problem	...	7
4	Present System of Examination	...	11
5	Some Educational Problems of Examination Reform	...	18
6	Some Technical Problems	...	31
7	Summary and Resulting Recommendations	...	39

APPENDICES

I	Comments of the Vice-Chancellor's Conference on the First Draft Report of the U.G.C. Committee on Examination Reform	...	46
II	Pass Percentage in I.A., B.A., I.Sc., & B.Sc. Examinations	...	49
	Pass Percentage in I.Com. & B. Com. Examinations	...	50
III	Percentage of Failures in M.A. Examination	...	51
	Percentage of Failures in M.Sc. Examination	...	52
	Analysis of Examination Results	...	52
IV	Questionnaire of the University Grants Commission	...	53

V	Tabular Presentation of the Questionnaire Study	...	56
VI	A Review of Admission Procedures Used in Colleges and Universities Abroad	...	65
VII	The Tutorial System	...	76
VIII	Grading Systems in Colleges and Universities of the United States and their Practice in India	...	78
IX	Grading System in Japan	...	86
X	The Combination of Internal Assessments with Examination Results	...	89
XI	The Marks of Examiners	...	97
XII	Classifying Students by Examination	...	102
XIII	Choice and Chance in Conventional Examinations	...	107
XIV	Evolution of Essay-Type Answers	...	112

P r e f a c e

There is widespread interest in the question of the reform of our examination system. It is generally recognised that there are serious disadvantages in the present system and that it is necessary to make the process of testing students relevant to the study undertaken by them and as accurate as can be made.

It would, however, be wrong to think that when we talk of reform, we are thinking of changing over from one set of practices to another set of rigid practices. Methods of testing should be related to practices of teaching and the conditions of study, and should be capable of adjustment and adaptation to particular situations and needs. Equally, however, when we talk about a system prevalent throughout a country, it is necessary to prevent haphazard changes taking place that would lead to confusion.

It is necessary first of all to identify the problem; then, there should be a fair measure of agreement on what is aimed at and the steps by which the desired result can be obtained. This Report of the Examination Committee appointed by the University Grants Commission is intended to draw attention to the problem and to indicate certain steps that can be taken to improve the methods of examination in our universities. Some of the general statements made in the report apply to the situation as a whole, but would not be applicable in detail to every particular course of study or department of teaching. For instance, the phrase "essay type of examination" would be more relevant to examinations in the Humanities and Social Sciences than to the Natural Sciences, even at the under-graduate stage. The problem of objectivity is much more difficult in dealing with subjects in which opinion and judgment are tested as much as information and possession of facts.

Persons engaged in the business of teaching and examining would do well to give some thought to this whole question and it is our

hope that this report will act as a stimulus to thought and possibly action. The present report has underlined the fact that "teaching, learning and examination actually constitute a unity of functions". Although the report has dealt at some length with the mechanical and technical aspects of examination, it has also emphasised the fact that any reform in examination would be unavailing without a corresponding reform in the teaching and learning processes. This report, together with an earlier document published by the University Grants Commission under the title "Evaluation in Higher Education", will, I hope, provide a comprehensive body of facts and suggestions that might lead to the solution of some of the current problems of the examination system in India.

On behalf of the University Grants Commission, I should like to thank the members of this Committee who gave a great deal of time and thought to the study of this problem and the production of this report.

New Delhi,
29th August, 1962

SAMUEL MATHAI
Secretary,
University Grants Commission

Introduction

It is generally admitted that the present examination system in Indian Universities needs reform. The system has been under criticism for a long time. While the need for reforming examinations has been recognised, and much discussion has taken place, little has been done so far to bring about improvements. Psychological resistance to changing the old system, inadequate financial resources to meet the cost of improvements, shortage of trained personnel for carrying out the measures of reform in a scientific way and the pressures to which the universities are subjected in favour of continuing the existing arrangements have all contributed to this. There is also a genuine fear in some quarters that the application of some of the proposed 'remedies' may possibly create a worse situation. But even those who have anxieties of this kind agree that continued inaction is bound to affect adversely the academic standards of the universities.

The most compelling reason for examination reform comes from academic considerations. It is not adequately recognised that examination is an aspect of the educational process which is intimately linked with its other important aspects—teaching and learning—and that teaching, learning and examination actually constitute a unity of functions. Teaching as well as learning are bound to be affected by a defective examination system, since both are dominated by the objectives that govern examinations. If examinations are, as at present, mainly a test of memory and do not satisfactorily measure the critical power and intellectual development of students, it is not surprising if students and teachers do not approach their work in the right spirit. Learning will then tend to become a mechanical acquisition of certain types of skills or information and teaching will become simply "coaching for

examinations". It can at once be seen that the relation between examinations on the one hand and standards of teaching and learning on the other is intimate. Improvements in any one of these areas involve a corresponding improvement in the other areas also.

In other words, examination reform has to be approached not merely from the narrow angle of improvement of evaluation techniques but from the larger stand-point of making our educational system better able to perform its function of developing the students. This would mean that we have to deal not only with the processes and procedures associated with examinations, but concern ourselves also with the broader issues involved in making our universities centres of good education. The problem here is how examinations can be made an ally of good education. When we approach the examination system in this way we are led to consider its mechanics as well as such matters as proper selection of students for university courses, clarification of the general and specific objectives of teaching, use of better methods of instruction, improvement of libraries, provision of residential facilities for students etc., which have a bearing on standards of education.

In the present context of the social and economic life of our country, examination reform has gained a new urgency. As pointed out by Dr. C. D. Deshmukh, former Chairman of the University Grants Commission, on the occasion of the inauguration of the seminar on 'Reform of the Examination System' on 10th July, 1958, at the Osmania University, "There have taken place in recent times in India great changes in the political, social and economic life of the country. We stand at a significant point in the socio-economic development of Indian society. In this context purposeful planning has become imperative in every important sphere of national activity. Our educational techniques and purposes have also in the circumstances to be adjusted to meet the new requirements of individual and national development. The examination of the educational system is, therefore, part of a process of earnest assessment of national goals and targets of achievement. It would be necessary, in other words, to reconsider the purposes of education and clarify the objectives to be aimed at".

Out of this evaluation has emerged the recognition of the interrelationships of educational processes — teaching, learning and examining — and the possibility of making examinations a potent instrument of educational reform.

Appointment of the Committee

The University Grants Commission is required by the U.G.C. Act, 1956 "to take, in consultation with the universities or other bodies concerned, all such steps as it may think fit for the promotion and co-ordination of university education and for the determination and maintenance of standards of teaching, examination and research in universities". It is thus one of the statutory functions of the Commission to consider problems relating to examinations and to bring about necessary improvements.

Already, in April, 1955, the Commission had drawn the attention of the universities to the desirability of reducing the burden of the final examination, distributing the work of the students uniformly throughout the course, instituting tutorial work, and trying out experiments of using objective tests and improving examinations. The Commission also requested Dr. Zakir Hussain, then Vice-Chancellor of the Aligarh Muslim University, to initiate a study of the question of improving the system of examinations and to prepare a pilot project in relation to the Aligarh University.

The preliminary measures referred to above were followed in September, 1957 by the appointment by the University Grants Commission of a Committee consisting of the following persons to advise the Commission on the steps that may be taken in this behalf:

- (a) Shri S. R. Dongerkery, Rector,
Bombay University, Bombay (Now
Vice-Chancellor, Marathwada Uni-
versity, Aurangabad) ... Chairman

- (b) Dr. Salamat Ullah, Principal Teachers' College, Jamia Millia, Delhi. ... Member
- (c) Prof. T. K. N. Menon, Dean of the Faculty of Education and Psychology, Baroda University, Baroda. "
- (d) Dean F. C. Ward, Educational Consultant, Ford Foundation, New Delhi. "
- (e) Dr. E. A. Pires, Principal, Central Institute of Education, Delhi. "
- (f) Prof. S. Mathai, Secretary, University Grants Commission, New Delhi. '
- (g) Dr. P. J. Philip, Development Officer, University Grants Commission, New Delhi. ... Member-Secretary.

Dr. H. J. Taylor, Principal, Scottish Church College, Calcutta and Shri D. C. Pavate, Vice-Chancellor, Karnatak University, were later added to the Committee.

The Committee first met on the 25th of October, 1957, at Delhi and prepared a rough scheme for examining the problem and also drew up a questionnaire to be addressed to the universities regarding the present objectives governing both teaching and examination in different subjects. The second meeting was held at Hyderabad on the 10th July, 1958, when the members, including Shri D. C. Pavate and Dr. H. J. Taylor, had an opportunity of discussing the issues involved with Dr. B. S. Bloom, Professor of Education and Head of the Department of Examinations, Chicago University. Dr. Bloom had been invited by the Education Ministry of the Government of India to advise on Secondary School Examinations and had found some time, through the courtesy of the Ministry, to assist the University Grants Commission to hold some seminars on the reform of the examination system. The Committee also participated in the seminar held at the Osmania University from the 10th to 13th July, on the objectives of teaching and examination

in Political Science and Chemistry. The third meeting of the Committee took place on the 5th September, 1958 at Delhi, when Dr. Bloom spoke about his impressions of the seminars and his subsequent meetings with student groups at university centres and there was also a discussion on the suggestions made by Dr. Bloom for reform of the examination system. At the fourth meeting, which took place on the 5th and 6th December, 1958, the Committee considered in a general way the immediate steps that should be taken by the Commission and the universities to bring about improvements in examinations. Dean Ward ceased to be a member of the Committee immediately thereafter on being transferred to the Middle East to take up a new assignment of the Ford Foundation. The Committee's fifth meeting was held on the 31st August and 1st September, 1959 for an informal exchange of views between the Committee and a team of Evaluation Consultants, who had visited certain university centres in India under the India Wheat Loan Educational Exchange Programme. The Committee met again on the 4th and 5th December, 1959 to consider the draft report. The report was also placed before a conference of Vice-Chancellors organised by the Ministry of Education at Khadakvasala on 15/16th June, 1960. Most of the recommendations made by the Committee were endorsed by this conference whose comments on the subject are contained in Appendix I.

Statement of the Problem

Successive Commissions on university and school education have pointed out the defects and weaknesses inherent in the examination system and suggested remedies of one kind or another. The University Education Commission of 1948 had remarked: "We are convinced that if we are to suggest one single reform in University Education, it should be that of the examinations". Among the more flagrant deficiencies singled out by the Commission were the following:

- (a) The chief purpose of the present examinations is not organically related to the actual process of education. Tests and examinations should be designed chiefly with educational ends in view.
- (b) The essay type of examination which prevails at the degree stage as elsewhere suffers from such major defects as lack of clearly defined objectives, arbitrary sampling and subjective scoring.
- (c) As no credit is given for class work, students do not apply themselves regularly to their studies.
- (d) Students experience considerable mental strain under the present system under which a final examination is held at the end of the degree stage on the result of which the future of the students depend.
- (e) The standards for success in examination differ in different universities.

During the twelve years which have elapsed since these comments were made, the scale of the problem has increased vastly.

The number of universities in India has more than doubled in the last two decades. The enrolment of students in colleges and universities has also registered a considerable increase over the years. Against about 4 lakhs of students in 1950-51, we have today nearly 10 lakhs in the various colleges and universities in India. This steep rise in the number of students over the past decade has increased the magnitude of the problem of examinations in India, for examination work is exacting and needs special skills.

It has been pointed out that our failure rate is perhaps the highest in the world today, often reaching 50 to 60 per cent in individual examinations. This is shown in Table I (Appendix II). It will be observed that the failure rate during the past 10 years has been roughly 50 per cent at the B.A., B.Sc., and B.Com. examinations. There can be no doubt that this high failure rate means that a large part of the expenditure on university and college education in India is either misused or lost. We seem to have resigned ourselves to this wastage and taken it for granted in most calculations on educational planning.

The high rate of failures at all public examinations upto the B.A., B.Sc., B.Com. levels, however, stands in contrast to the performance of students in the M.A. and M.Sc. examinations. The data available for the year 1958-59, reveal that at the M.A. and M.Sc. examinations the failure rate was as low as 18.6 per cent and 19.3 per cent respectively. Indeed 11 of the reporting universities had a pass rate of 90 per cent or above at the M.A. examinations. The same trend is evident in the case of the M.Sc. results with the difference that the combined percentage of 1st and 2nd class students at the M.Sc. examination is better than the corresponding ratio on the Arts side (see Appendix III). The comparatively low rate of failure in the post-graduate examinations as against the high rate of failure in the first degree examinations, which has been a consistent feature of university examinations, particularly during the last ten years, is significant.

The qualitative and quantitative superiority of the results at the post-graduate level appears to be due to greater care in the selection of students for the various courses and provision of better facilities

for teaching and learning. By and large, the quality of students who are admitted to the post-graduate courses is superior to that of the under-graduates. There is also a greater emphasis at the post-graduate level on the aptitude of students for the subjects in which they seek to specialise. Again, the facilities that are provided in the universities for post-graduate instruction are generally superior to those available for under-graduate teaching. Greater stress is also laid on written work at the post-graduate level in the form of theses and essays for Arts students and 'practicals' for Science students. This makes the student more regular and painstaking in his studies.

The results of post-graduate examinations thus underline the conclusion that any worth-while reform in the examination system involves not merely an improvement in its techniques, as is commonly supposed, but also betterment of the conditions in which the teaching-learning process is carried on in universities and colleges in India.

It is not sufficient to condemn the present system of examination in general terms. We need a more detailed analysis of the educational system and the conditions under which it actually operates. Improvements can only be made on the basis of knowledge. Some enquiries to this end have been undertaken in a number of places, for example, at the universities of Aligarh, Baroda and Calcutta, the Indian Statistical Institute, Calcutta, the Institute of Technology, Kharagpur and the Bureau of Educational Research, Ewing Christian College, Allahabad. While such studies are of great value, we need a continuous stream of applied research of the highest quality in many more centres to reveal the full range and implications of the problems we have to solve.

We therefore feel that one of the most urgent needs in regard to examination reform relates to study of features which contribute to the defects of the examination system. These may include enquiries regarding both the educational and technical aspects of examinations. The educational problems to be studied may include: (a) different kinds of requirements for entrance to universities, (b) aptitude tests for university courses, (c) sociological

factors influencing student performances, (d) suitability of various types of examinations, e.g., objective tests, open book tests, multiple choice answers, short answer tests, etc. for evaluating student achievement, (e) teaching methods in relation to performance in examinations, etc.

The following questions raised by Dr. Bloom pertaining to the soundness of the present examination procedures in Indian Universities also deserve study :

- (1) What is the level of reliability of these examinations and by what techniques can it be improved?
- (2) What is the present level of error attributable to subjectivity on the part of the examiner and by what technique can this error be reduced?
- (3) What are the soundest procedures for combining marks from internal and external assessments?
- (4) What is the relationship between the student's performances in various papers and how should the results be combined?
- (5) What is the relationship between the student's examination marks and the teacher's judgment of his competence?
- (6) What kind of examination questions do the students elect to answer and what is the nature of the questions they avoid?

Research along these lines may be undertaken in the universities with the help of their departments of Education. Perhaps the newly constituted National Council for Educational Research and Training may also be able to take up such investigations. It will, however, be necessary to disseminate the findings of these enquiries to the universities for their information and guidance.

Present System of Examination

To elicit information about the prevailing system of university examinations, the University Grants Commission circulated among the Indian Universities a questionnaire (see Appendix IV) consisting of 21 questions under seven broad heads. In response to this, 38 universities sent their replies along with relevant papers and documents regarding their respective systems of examination. An analysis of the data thus obtained is given in Appendix V.

A study of these data reveals that there is a great diversity in (i) the nature of examinations, (ii) the scope and application of internal assessment, (iii) the nature of question papers, (iv) the marking of examination scripts, (v) the classification of examination results, (vi) research in examinations and (vii) the contemplated reform in the system of examination. Following are the main findings:

Nature of Examinations

By and large the written essay type examination dominates the present system of examinations in India. An oral test (*viva-voce*) forms part of the practical examinations in science subjects in all the universities at all levels. A *viva-voce* also forms part of the examination of a thesis when it is submitted either at the post-graduate or at the doctoral level. But in no university does the *viva-voce* have the same status as the written or practical test, in the sense that success or failure in the examination is mainly determined by the candidates' performance in the *viva-voce* test. Medical examinations are perhaps the only exception to this general rule.

In some universities sessional work forms a part of some

examinations, which means that evaluation of the sessional work does contribute to the final evaluation of the candidates' performance in these examinations. In certain engineering examinations sessional work has a distinct status like that of any written examination paper.

Internal Assessment

The value attached to the assessment of the day-to-day and periodical work of the students varies from one university to another. Whereas all universities make internal assessment an organic part of the final assessment of candidates' performance in some examinations (particularly science, technical and professional ones), there is a wide variation in the marks allotted for internal work in different examinations in the universities. Marks allotted to internal work in some of the subjects vary from 5 per cent to 100 per cent, the maximum of 100 per cent being the case in certain practical aspects of engineering courses. Sessional work comes into the picture chiefly in science and engineering examinations and in certain professional examinations such as those in Education, Law and Medicine.

Information about the contribution of internal assessment to the final aggregate score is very scanty. Only 8 of the 38 universities indicate that marks allotted to sessional work in the B.Ed./B.T. examinations form a part of the candidates' aggregate score. No corresponding data are available from the other 30 universities. In the case of Jadavpur University, the marks awarded for every day work are taken into consideration for determining honours in the final examination, for the degree of Bachelor of Engineering.

Visva-Bharati reports that the assessment of sessional work which had been introduced as a part of the final examination has been dropped in the School Final Examination and it is proposed to drop it in the higher examinations as the system has been found to be unsatisfactory. Lucknow also has reported that in Arts subjects the system of assessing sessional work as part of the final examination has not produced satisfactory results owing to lack of suitable teaching staff. It would thus appear that different

universities have different attitudes towards internal assessment and have responded to the idea differently. The other eleven universities have found internal assessment helpful in properly evaluating candidates' work.

Question Papers and their Moderation

So far as paper-setting work is concerned, 11 universities (Annamalai, Baroda, Bombay, Gujarat, I.A.R.I. Delhi, Karnatak, Kerala, Marathwada, Poona, S.N.D.T. Women's and Sri Venkateswara) adopt the practice of paper-setting by a Board of Examiners at both the post-graduate and under-graduate levels. In Gujarat University there is no fixed rule. Generally the paper-setting in this university is done by two examiners jointly. In the Andhra University the paper-setting work in under-graduate examinations is done by a Board of Examiners, but in the post-graduate examinations the university has one examiner for each paper. As many as 22 universities (Agra, Aligarh, Allahabad, Banaras, Delhi, Gauhati, Gorakhpur, Jadavpur, Jammu & Kashmir, Kuruksetra, Lucknow, Madras, Nagpur, Osmania, Patna, Rajasthan, S. V. Vidyapeeth, Saugar, Utkal, Varanaseya Sanskrit, Vikram and I.I.S. Bangalore) have one examiner for each paper both at post-graduate and under-graduate levels. Panjab university has two examiners at the post-graduate level and also for the B.A./B.Sc. Honours papers. The other 3 universities (Calcutta, Roorkee and Visva-Bharati) have more than one examiner for the post-graduate examinations and only one at the under-graduate level. In professional examinations like medicine and engineering there is no uniformity of practice regarding the number of paper-setters to be appointed.

From the data supplied by the universities it is found that 13 universities (Annamalai, Baroda, Bombay, Gujarat, Jammu & Kashmir, Kuruksetra, Marathwada, Panjab, Poona, S.N.D.T. Women's, Vallabhbbhai Vidyapeeth, I.I.S. Bangalore, I.A.R.I. Delhi) do not moderate their question papers at all.

The principal reason for moderation (in these universities where questions are moderated) is almost the same in all cases, viz., for preventing the questions set from going beyond the boundary of

the prescribed syllabus. The moderator also sees that the questions do not overlap even within the syllabus and that an overall uniformity of standard is maintained.

Saugar University has stated that the work of moderation aims at seeing whether the instructions given to the paper-setter have been strictly followed, particularly those in regard to the distribution of the questions over the subject area and the allotment of marks. Osmania and Banaras have stated that moderation aims at correcting the ambiguities and obscurities in the language of question papers.

As has already been noted above, some 13 universities do not moderate their question papers and have therefore no machinery for moderation. Twenty-one of the other universities have a Board of Moderators consisting of more than 2 or 3 members. In most cases the head of the department is a member or the chairman of the Board.

Jadavpur, Karnatak and Sri Venkateswara report that question papers are moderated when necessary, but they have no separate machinery for moderation. Calcutta and Gauhati have the question papers moderated, not by a Board, but with the help of more than one expert.

Marking of Examination Scripts

Andhra, Annamalai, Baroda, Calcutta, Gujarat, Jadavpur, Karnatak, Kerala, Madras, Osmania, Roorkee, Panjab, Sri Venkateswara and Vallabhbhai Vidyapeeth have the system of appointing two examiners for each "paper" at the post-graduate level and sometimes at Honours examinations. Usually, the mean of the awards is taken as the final mark for an answer paper when the difference is below 10 per cent. Roorkee appoints a third examiner when such difference is 20 per cent or more, Rajasthan, 15 per cent or more, Jadavpur, 12 per cent or more and Madras, more than 10 per cent. Four universities (Allahabad, Lucknow, Visva-Bharati and Rajasthan) appoint more than one examiner for the thesis papers of

post-graduate candidates only. In Utkal University more than two examiners are appointed for the M.B.B.S. examination.

In most of the universities no scaling of marks is done after the examiners submit the mark sheets. Any "moderating" of marks is not scaling in the true sense of the term. While 23 universities have no scaling of marks, in 14 of them some moderation of marks is attempted at the examiners' meeting with the head examiner or paper-setter or head of the department.

To the question whether examiners are given any instructions regarding the spread of marks corresponding to different grades of answers by the candidates, 21 universities have reported that no such instructions are given. While 3 universities have stated that such instructions are issued, there is no indication about the nature of the instructions.

The number of examination scripts given to a single examiner varies not only from university to university but also from one examination to another, the range of variation being 10 to 500. Barring the extreme cases, it may be said that the average number of answer scripts given to a single examiner lies between 200 and 300. Eight universities give as many as 500 scripts to an examiner. The University of Madras reports that the allotment of examination scripts depends on the time available to the examiners.

Classification of Results

In Honours examinations almost all universities classify successful candidates into First Class and Second Class only. In the M.A./M.Sc. and professional examinations of as many as 15 universities (Agra, Allahabad, Annamalai, Gauhati, Gorakhpur, Jammu & Kashmir, Kuruksetra, Lucknow, Madras, Panjab, Patna, Rajasthan, Saugar, Utkal and Visva-Bharati) there are three classes. In some universities such as Karnatak, Banaras, Vallabhbbhai Vidya-peeth, and Sri Venkateswara, there is another class of successful candidates who are termed First Class with distinction, I.A.R.I. Delhi places all successful candidates in classes A, B and C.

Most of the universities place in the First Class/Division, those candidates who secure 60 per cent of the marks or above in the aggregate. For a Second Class/Division, the general rule appears to be that a candidate should secure 50 to 59 per cent of the marks in the aggregate. As regards the minimum percentage of marks for a Third Class/Division, there is considerable variation between one university and another. Most of the universities have fixed the minimum percentage for a Third Class/Division at 33, 35 or 36. Delhi and Jammu & Kashmir have not mentioned the minimum marks required for a Third Class/Division. Ten universities (Annamalai, Baroda, Bombay, Kuruksetra, Madras, Panjab, Rajasthan, S.N.D.T. Women's, Sri Venkateswara and Visva-Bharati) have fixed 40 per cent of the aggregate for a pass in the Third Class/Division. Sri Venkateswara, Madras and Banaras state that 75 per cent of the marks are necessary for a First Class with distinction, for which Vallabhbhai Vidyapeeth has fixed 66 per cent marks in the aggregate. Roorkee confers Honours on a candidate who secures 75 per cent of the marks. I.A.R.I., Delhi classifies students into classes A, B, C and D which are not based on numerical marking. In Utkal, I.I.S., Bangalore, and Rajasthan, the percentage of marks required for different classes of success varies from examination to examination.

Research and Reform in Examinations

It is very significant that 31 of the 38 reporting universities have categorically stated that they have neither undertaken any research work in examinations, nor do they possess any psychometric units.

It is also remarkable that 19 of the 38 reporting universities have not introduced any reform in their system of examinations.

Of the remaining 19 universities as many as 13 have started taking into account sessional work in the final ranking of candidates besides their performance in the external examination. They are Aligarh, Baroda, Delhi, Jammu & Kashmir, Karnatak, Kerala, Lucknow, Panjab, Patna, Saugar, Vallabhbhai Vidyapeeth, Visva-Bharati and I.I.S. Bangalore. The term "sessional work" includes

tutorial work, laboratory practicals, seminar work and internal tests and examinations.

Some universities have introduced a novel system to prevent malpractices in examinations. They transform the actual roll number of candidates into different code series before the scripts are given to examiners for assessment. This is intended to prevent the bias of the examiner from coming into play in the evaluation of a candidate's performance.

In many universities degree examinations are held in parts instead of as a single whole at the end of the course. A *vive-voce* test is obligatory in post-graduate examinations in Baroda and in the Acharya (Final Sanskrit) examinations in Varanaseya Sanskrit Vishvavidyalaya. Only Baroda University has so far introduced objective type of questions and short answer questions.

The reforms that may be brought about in the examination system may be divided into two categories — reforms which are, by and large, of an educational nature and reforms which are linked up with the technical aspects of examinations. To the former belong problems that relate to selection of students for various university courses, internal assessment, objectives of teaching different subjects etc.; to the latter such questions as appointment of examiners, marking of answer scripts, scaling of marks, etc. We shall take up the educational problems first.

Some Educational Problems of Examination Reform

The problem of selection

We have already called attention to the high rate of failure in university examinations. To some extent this may be attributed to the vagaries of marking, and to the insistence on separate "pass marks" in a series of subjects. Another contributory factor is the inefficient teaching and inadequate facilities to be found in many of the colleges. But we believe that one of the most important causes of the high failure rate is that universities admit many students who are intellectually incapable of the work required of them, or who are not sufficiently interested in learning. It is not enough to make improvements in the educational system; we must also ensure that the right persons are admitted to it. The selection of students is central in any discussion of educational reform, and in our view great stress should be laid on this question.

As the Report of the U.G.C. for the year 1959-60 points out, our good students are as good as many students in any part of the world. But we have in our universities far too many unfit students who have come in merely because they did not know what else to do and because no stiff standards were applied in admitting them to the university. Enough study has been carried out to show that a good many of the students who actually enter the universities and many who seek to enter the universities do not in fact have the necessary intellectual and emotional preparation for higher education. It is not fair to thousands of young men and women to place them and their parents in a situation in which they have no alternative but to seek admission to universities.

A selective process of admission in universities and colleges is prevalent in most countries in the world today. In no country in the world is university education thought of as a universal right for which the State should make provision. Comparative statistics of other countries including the industrially advanced countries of the West indicate that India has a relatively high proportion of university students in relation to the number of young people who go to school and complete their school education. It is also necessary to view the problem of admission to colleges and universities in terms of academic and national needs and in relation to the availability of resources.

The present method of choosing students for admission may be more or less stringent, according to the accepted standing of the institution which applies it. But it does not operate to any great extent as a true selection for admission to university studies, at least in those universities which affiliate a large number of colleges. The process merely distributes the applicants in such a way that the more reputed colleges secure the best students, and the less reputed colleges take the remainder. This process tends to perpetuate the differences between the colleges, a state of affairs which is in many ways undesirable.

Our recommendation with regard to a selective process of admission in the universities and colleges does not imply the application of a battery of ability and aptitude tests to each student who seeks admission to a university course. For one thing, it is doubtful whether such precise testing is necessary or desirable or even feasible in the case of those who wish to have a general education; for another, even if it were so, we do not have enough personnel and resources to undertake such a gigantic task. We do, however, want to stress that no reform in the system of examination will reduce the failure rate in our universities and colleges unless the prevailing admission procedures are improved. The problem is really two-fold: first, we have to see that only those candidates are admitted to colleges who are fit to profit by higher education; and secondly, we have to make sure that our resources are not wasted in fixing square pegs in round holes.

In England the decision as to what broad courses a student will take in his college is made at the age of 11 plus by holding a special examination for the purpose. The admission procedures used in the U.K. and other countries abroad are described and analysed in the note contributed to this report by Dr. Rhea S. Das of the Indian Statistical Institute of Calcutta (see Appendix VI). While for a variety of reasons we may not be able to adopt many of these procedures in India, ways have to be found for determining the suitability of candidates for university education.

Various suggestions have been made with regard to the means that may be adopted for a more careful selection of students for admission to the universities. One of these is that special tests may be held by the universities and/or colleges as is done now by some professional institutions. The University Grants Commission considered this suggestion but felt that it was not practicable to arrange for such special tests in view of the large numbers involved and because of the time factor. Something like six lakhs of students now pass the School Leaving Examinations and nearly 55 per cent of them get admission to universities. Obviously, the number of those who seek admissions is very much larger than the number admitted; it may be in the neighbourhood of five lakhs. To hold special examinations for such a large number would be an expensive operation. Furthermore, the results of the School Leaving Examinations are announced only a few weeks before the opening of colleges and universities; and to ask students who have just gone through a strenuous public examination to prepare themselves for another test seems hard and impracticable.

Perhaps one solution of the problem may lie in the direction of basing the admissions to the university on the cumulative records of students at the school, the maintenance of which was advocated by the Secondary Education Commission. These records may be utilised for finding out whether a particular student is suited for university education and, if so, what is the broad area in which he should conduct his studies. This of course means that our secondary school teaching and examination will have to be appropriately reorganised.

Yet another suggestion is that students might be interviewed and a *viva-voce* test given to them before admission to the universities. Some institutions already have an arrangement of this kind, but to adopt this suggestion as a universal rule is not easy, as it would mean that students would have to travel from their homes to the colleges and universities for this *viva-voce* test and for the interview; and if admission is not secured they would incur considerable expenditure fruitlessly.

A suggestion worth considering is introducing in the School Leaving Examination one or two additional papers for those who wish to compete for entrance to universities. These papers may be tests in the student's mastery of the language to be used at the university and of his intellectual maturity and general interests. Students who do well in this part of the test and at the same time secure qualifying marks in the general examination may prove to be more suitable material for admission to the university; and selection from amongst them in order of merit may prove a satisfactory way of selecting students for university education.

The proper place of examinations in education

The University Education Commission pointed out that examinations bear very little relation to the wider purposes of the educational process, but they did not discuss the question in detail. We agree that this is a fundamental weakness. It is fully recognised that reform of the examination system involves finding satisfactory answers to certain specific problems, such as the selection of examiners, the spacing of examinations, the introduction of more objective marking, the tabulation and combination of marks, etc. But reform does not consist wholly in finding the answers to these problems. It is necessary to put the examination in its proper setting as something which contributes to the total purpose of education. It must not be regarded as an end in itself and allowed to dominate, as it so largely does at present, the whole educational system. Teaching, learning, and examining constitute an indissoluble trinity of functions in an academic community, and they must be related to each other, and to the objectives which the educational process is desired to achieve.

It is, therefore, necessary to state, and continually re-state, these educational objectives. This can best be done by groups of teachers, in seminars or 'workshops' of the kind recently conducted by Dr. Bloom. There are certain general objectives which would be valid for any liberal education, and there are the particular objectives of the different subjects. When the objectives are framed, it is possible to formulate more clearly the specific nature of the knowledge and competence which students are required to demonstrate in an examination. Such definition helps to draw up a syllabus, which while indicating the range of material to be studied does not become a mere catalogue of topics. A syllabus properly devised should state not only the minimum requirements in a subject but also indicate wider possibilities of study and learning, and thus encourage independent thinking and study on the part of the student. It is also necessary continually to revise the syllabi in accordance with the living developments in the fields of knowledge concerned. We are glad to note, in this connection, that the U.G.C. has recently appointed 'Review Committees' of experts in different Science and Arts subjects to examine the content of syllabuses in the universities and to define in a clear way the objectives to be aimed at in the teaching and learning processes at different levels. We also welcome the proposal of the Commission to hold summer schools and seminars in various subjects at different university centres, as such meetings of teachers will help to clarify the goals of teaching in those subjects.

Unfortunately the setting of a question paper has come to be recognised as an annual ritual in which the examiner puts out questions, with slight verbal changes from the sets of question papers of previous years. The skill of the examiner is often measured by his adroitness in defeating the "guess" of the students and their teachers. This game of guessing determines the way in which students prepare for the examination and leaves them open to the temptation of depending on the "guess" of professional guess-makers and writers of "made-easy" notes. The broader aims of education and the specific objectives of teaching thus get lost.

There is nothing revolutionary in the suggestion that a clear definition of the broad aims of education and of the specific

objectives of teaching and examination in different subjects at different levels is essential to the success of any teaching and learning process. In fact a necessary condition for the success of any serious activity, it will be admitted, is a clear understanding of the goals to be pursued by it. It is not enough to formulate the broad aims of education; more important, even though a little more difficult, is to define these aims in terms of the requirements of particular subjects of study. Our teachers and students are at present largely dependent on syllabuses which catalogue the content of courses and recommend books without stating clearly the specific objectives which should inform the teaching and learning activity in any particular subject. Teachers and students are left free to interpret the syllabus as best they can and are often tempted to take the path of least resistance which is to concern themselves with just the minimum which may ensure a reasonable chance of success in the examination.

To think out and to state clearly the objectives of education in each field is a necessary condition of any real advance. This cannot be too strongly emphasised, and yet such thinking seldom takes place in the universities. Syllabuses have been handed down from year to year and from decade to decade without much change, with an almost exclusive emphasis on information. Was it not T. S. Eliot who asked: "Where is the wisdom we have lost in knowledge? Where is the knowledge we have lost in information?" This undue emphasis on factual information has encouraged students to memorise without understanding and has prompted the widespread use of cram books. It has also been responsible for the use, in most examinations, of a stereo-typed kind of question paper consisting largely of questions to which set answers can be given.

Relation between teaching and examination

It cannot be stressed too much that improvement of examinations also entails a parallel improvement in the processes of teaching and learning, for each reacts upon the other. We must get rid of the idea that the teacher's business is limited to the giving of a certain number of lectures. If the quality of teaching is to be improved, more time must be devoted to tutorials and seminars, which bring

the student more intimately into contact with the mind of the teacher. If this were done, the number of lectures could in many cases be reduced with advantage. The information collected in regard to teacher-student ratios, and the physical and other facilities available in universities and colleges shows that there is much room for improvement. Deficiencies will have to be rectified by increasing the number of teachers, by making personal contact possible, and by providing more accommodation for both teachers and students. It will also be necessary to improve college and university libraries by the addition of a number of copies of standard books and the provision of additional sitting accommodation to encourage students to cultivate the habit of independent study and wide reading.

The value and significance of the tutorial system as an ally of good education in the universities have been widely recognised and in India its importance has been emphasised by the University Education Commission.

It may at once be admitted that it will be difficult to introduce in Indian universities the system of tutorials which prevails in Oxford and Cambridge, because of the paucity of personnel and accommodation, from which most of them suffer, and the fact that they are largely non-residential institutions. The colleges in Oxford and Cambridge are primarily residential units where the teachers and students live together. Students not living in colleges live in recognised lodgings within easy access of the colleges. At Oxford each student gets at least one hour's tutorial instruction a week. A tutorial consists of reading of an essay on an assigned subject followed by a discussion, each tutor meeting one or two pupils at a time. The 'supervision' system at Cambridge is also similarly organised.

In the provincial universities of England, where the time at the disposal of the students and teachers is limited (because many of them live far away from the university campus), a system of tuition in small groups is followed. Members of the group read short essays in rotation week by week. The tutor then stimulates

discussion and comments more fully than he can in his lectures on particular points which need elucidation or development.

In the London School of Economics and Political Science there prevails what is known as the "class system". All the members of a "class" are expected to have acquired an appropriate background in the subject to be discussed. Either one member reads a short paper on the theme or the tutor himself gives a short introduction to the subject under discussion. In the latter case the "class" or the tutorial becomes a substitute for a lecture.

In the American universities and colleges more importance is attached to the discussion than to the formal lecture as a method of instruction. The students are placed in discussion groups of 10 to 15. Two or three days before a topic is taken up for discussion, the instructor gives reading assignments to his class. The students are expected to read the books and other materials prescribed before coming to the discussion, so that they are in a position to answer questions on them and take an intelligent part in the discussion. Even in the regular classes questions by the pupils are encouraged.

From the foregoing it will be clear that the system similar to that adopted by the provincial universities in the U.K. or by the colleges in the U.S.A., is more suitable to Indian conditions than the Oxford and Cambridge system. The tutorial group in our universities and colleges may have to consist of 10 to 15 students. While the various details of organising tutorial work are best left to the universities themselves (to be planned preferably at periodical seminars or conferences of teachers), two things are essential, viz., (i) writing of essays by students and (ii) discussion in the class of the points considered in the essays, the students being required to take an active part in such discussion. (As an illustration the suggestions made by the Bombay University—by its former Rector—for holding tutorial classes are attached as Appendix VII).

Even to introduce a seminar system such as the one described above, the teacher-pupil ratio and accommodation in most of our universities and colleges will have to be improved. It is, however, possible to make better use of the existing staff and

buildings by cutting down the number of lectures at present provided by about 50 per cent. (This suggestion will not be applicable to courses which already provide for a sufficient number of tutorials or 'practicals'). The existing practice of "covering" the syllabus in lectures could be given up and lectures related to the salient aspects of a subject rather than to its details. The role of lectures should be conceived as supplementing the tutorial work rather than supplanting it. This does not, however, mean that lectures are unimportant. They have a useful part to play.

Teaching cannot be improved without able teachers. It will, therefore, also be necessary to attract and retain the services of such teachers in our universities and colleges. This problem will have to be tackled in an effective way if the right kind of teaching capable of inspiring students to approach learning in the right spirit is to be strengthened. To quote from the U.G.C. report for 1959-60: "the most essential quality required of a teacher is a sound knowledge of the subject which he undertakes to teach and the ability to communicate that knowledge to his pupils. A genuine interest in youth and an understanding of its psychology also play a part in the making of a good and successful teacher. An inquiring mind in the teacher serves as a spark to ignite the tinder of the mind of the young student who may thus develop a true love of learning and enquiry. In consequence of the shortage of adequately qualified men and women it will be necessary to attract and retain the services of such teachers in our universities and colleges. The universities have frequently to be satisfied with persons who have just the minimum qualifications. With the increase in the number of students and of subjects to be taught, the supply of good teachers has been steadily decreasing and the economic law of diminishing returns seems to be operating in the field of university education. The question of adequate supply of qualified teachers is indeed one of the most critical problems facing the universities today"

The importance of regular writing exercises by students cannot be too much stressed. Writing has in turn to be based on reading of books. It is, therefore, to be ensured that sufficient number of the requisite books (say 10 copies of each) are provided in the

college libraries. It will be necessary to improve university and college libraries by providing more books and more accommodation. A large number of colleges in the country do not at present have adequate libraries. We are glad to know in this connection that the U.G.C. has given a high priority to the development of university and college libraries and that already the Commission has disbursed considerable amounts of money for this purpose. It is, however, necessary to ensure that the facilities provided are fully made use of. It is perhaps not incorrect to say that the majority of our students have still to learn the art of reading and of spending time in a library profitably. A competent and imaginative teacher will know how to encourage the reading habit among his students. He will show them that insight in any subject comes from wide and selective reading and that "prescribed" text books often provide only the minimum material for study.

Internal Assessment

The mental strain involved in the present system of examinations has many evil effects, as the University Education Commission pointed out. Moreover, the whole purpose of the student tends to be narrowed down to the single end of acquiring a degree which he looks upon, as a distinguished educationist once put it, "as having an intrinsic rather than a symbolic value". The student confines his study to those portions of the course which are likely to be covered by the questions, memorises stock material, and makes little attempt to acquire a deeper understanding of his subject. Colleges, in their turn, become agencies for the supply of stereotyped information. They may instruct their students, but they fail, very largely, to educate them. These evils would undoubtedly be lessened if the dominance of the final examination were reduced.

For an examination to be a worth-while and educative process, it is necessary to emphasise that examinations should periodically reveal to the student his own progress in studies. Furthermore, for effective teaching it is necessary that the teacher should know at regular intervals how the students are progressing. The present system of one final examination at the end of the academic course

does not provide any inducement either to the teacher or to the taught to improve his work.

We do not suggest that the final examination should be abolished; this would be neither possible nor desirable. It is necessary to evaluate in some way the student's progress and development resulting from the educational process which he has undergone, and in the conditions of Indian Universities there is no acceptable substitute for examinations of some kind. But we do suggest that the nature of the final examination could be suitably modified. It need not necessarily take the form of a single concentrated performance on which the student's ranking, and therefore his future, solely depends. It should be possible to combine with the examination an internal assessment, based on a record of class work maintained by the teachers. The right type of internal assessment would stimulate the student to develop better methods of study, habits of continuous work, and the right attitude to learning. In the initial stages, doubtless, only a small weightage should be given to internal assessment, but if the system were wisely developed, the weightage could increase at a later stage.

We may refer here to the report of the University Grants Commission's Committee on General Education which emphasised the place that might be given to internal assessment in the teaching and examining of courses in General Education. The Committee felt that these courses would lend themselves easily and fruitfully to a system of internal assessment. Rightly developed, this would stimulate the student to choose reading materials with discrimination, to study the same with understanding and also help the student to gauge periodically the depth of understanding he has reached in a given course. While not advocating a total abolition of the final test, the Committee recommended that due weight (say 50 per cent of the credit in a given course) should be given to the assessment by the teacher in the class room. We, however, think that in the case of affiliating universities the scheme of internal assessment might first be tried in the university colleges which are situated in the university campus or near its headquarters, so that the university may supervise the methods of evaluation conveniently and at short intervals.

It should be however mentioned that the combination of an internal assessment with an examination result involves technical problems. The object of combining the two is to secure a more accurate measure of the student's attainment than is given by the examination result alone. But the combined result is not necessarily more accurate, and may indeed be less accurate, than the examination taken by itself. Whether it is more accurate or not depends on the relative reliability of the two results which are combined, and on the relative weight assigned to each. Progress has recently been made in the discussion of this problem, and the danger of assigning an *ad hoc* weight to the internal assessment has been pointed out. We call attention to the importance of this question, and to the need for further investigation. (A note by Dr. H. J. Taylor attached — Appendix X — may be seen regarding the problem of combining internal assessments with the examination result).

Some countries, notably Japan and the U.S.A., have generally favoured internal assessment by class teachers in preference to annual external examinations. Two notes on how the system of internal grading works in the U.S.A. (Appendix VIII) and in Japan (Appendix IX) are attached.

We are fully aware that there would be difficulties and dangers in introducing internal assessment in our universities. In some cases the award of credit for class work might be abused, and colleges might be exposed to the temptation of being too liberal in their assessments for the sake of producing impressive results. It has been reported that the Karnatak University which had allowed 20 per cent of marks for internal assessment had recently to abrogate this provision, because the assessment was "inadequately trustworthy". Clearly, there must be precautions and safeguards if the assessment is to be impartial and reliable. But the Committee feel that the educational merits of internal assessment outweigh the difficulties and that it is not beyond human ingenuity to find a satisfactory solution to this problem. Means can be found for preventing or minimising the influence of non-academic considerations in the internal assessment. It has been suggested, for example, that marked variations in the standard of internal assessment

could be detected (and corrected) by correlating the mean assessments of colleges with the performance of the same colleges in the external examination. A properly qualified officer of the university may also go round the affiliated colleges to check the reliability and validity of their methods of assessment. Madras and Rajasthan Universities have similar procedures for checking the internal awards of their affiliated teachers' training colleges.

We now turn to a consideration of some problems which are of a technical nature.

Some Technical Problems

The problem of marking

It has been universally recognised that one of the most difficult problems associated with the present system of examinations is the subjectivity involved in the marking of scripts. The early studies of the examination system initiated in the nineteen thirties in several countries, began with a survey of this problem. It was shown that the marks awarded separately by different examiners to the same script seldom agreed, and often differed very widely, even to the point where one examiner would 'fail' a candidate to whom another examiner would award a first class. Again it was found that when the same examiner revalued the same papers after an interval of time, the marks differed very considerably from those he gave on the first occasion.

Mention may be made in this connection of some case studies (Appendices XI, XII and XIII) made by Dr. H. J. Taylor which show that the present system of marking of answer scripts and scaling and combining marks, generally followed in our universities, lack reliability and validity. The present methods of marking examination scripts and of combining and tabulating marks in university examinations without reference to recognised statistical procedures are seen to be highly unsatisfactory. It is true that sometimes moderators are appointed to adjust the differences between the standards of the various examiners, but the methods they use often do not have the effect of bringing the marks to a common scale. The methods employed are usually not only inadequate, but sometimes wholly fallacious, and may even increase the discrepancies they are intended to remove.

It is important in this connection to consider the ways by which marks are combined into a final result. We may in fact distinguish five different methods of determining a student's result in an examination :

(1) According to the present method adopted in most university examinations, marks are awarded according to the judgment of the examiners and are simply added up to give the total.

(2) After the marks have been submitted by the different examiners, they are scaled by some suitable procedure in order to bring them to a common standard (defined by the mean and the standard deviation). The scaled marks are then added to obtain the total.

(3) The marks are used only for the purpose of ranking the students in the different subjects and the student's final result is determined by this rank in the whole group of candidates rather than by an absolute mark.

(4) Examiners are given detailed instructions regarding the marks they are to award to each type of answer, *e.g.* very good answers might be given 70 per cent marks and upwards, good answers from 50 per cent to 70 per cent, average answers from 30 per cent to 50 per cent and bad answers below 30 per cent. By defining the spread of marks in this way it is expected that the marks of different subjects will show much the same average and range. It will then be permissible to add up the marks and obtain the total without applying scaling techniques.

(5) Assessment is made in terms of grades rather than marks, a small number of grades only being recognised. Usually there are five (A) very good, (B) good, (C) fair, (D) poor and (E) bad. Weightage is given to grades obtained in the different subjects to obtain the final grades. This method is widely employed in America and elsewhere and has been used also in India.

We believe that method (1) has very little to commend it, for it is clear that the large inaccuracies involved are not removed to any great extent by the present methods of moderation which the

universities adopt. Methods 2, 3, 4 and 5 should be given serious study and should be tried out in practice. It would, however, be necessary for universities using these methods to do so with expert advice, to ensure that the conditions necessary for their validity are satisfied. We are informed that method (4) is employed in certain non-university examinations and that the Co-ordination Committee of the U.G.C. also has recommended it. This method has great merits in the absence of statistical methods for scaling the marks awarded by examiners.

The problem of scaling

The problem of scaling arises: (1) when the examination scripts are so numerous that they must be divided amongst many examiners: (2) when the marks in different papers and subjects have to be combined into a final result. We may consider these two cases separately.

In dividing scripts amongst examiners, it is important that the distribution should be random. Suppose that scripts from one centre go to examiner A and from another centre to examiner B. The mark-distributions of A and B may differ widely: but it is impossible to determine whether this difference is due to different standards of marking or to inherent differences between the two sets of scripts. If the scripts are mixed and then divided at random between A and B, we ensure that the two groups are statistically equivalent. The mean performance of the candidates, and its variance, will then be closely similar in the two groups. Differences in the mark-distributions can then be attributed to the examiners, and can be corrected by scaling. This is a technique by which the marks of one or both examiners are adjusted until the mean and the variance are approximately the same for both. We can then be sure that the work of all the candidates is being evaluated on the same standard.

In large examinations, the randomisation of scripts and the scaling of marks, even without introducing any changes, would result in a very great increase in the reliability of the results.

The problems are more difficult in scaling marks in different papers and subjects. The variance, or spread of marks, is usually very much higher in some subjects such as Mathematics than in others such as History and languages. It is the variance rather than the average mark which determines the weightage given to any particular subject when the marks of several subjects are combined. It would, therefore, be necessary to decide the relative importance to be attached to different subjects before an appropriate scaling procedure could be adopted. At present, when there is no scaling, some subjects are allowed to exercise a preponderating influence on the results of an examination even though the maximum marks assigned for the various subjects may be the same.

Objective tests

In certain countries particularly in the U.S.A., an attempt has been made to eliminate the subjective factor by replacing the essay type of examination by objective tests. In India, as stated earlier, the essay type still dominates the examination system and objective tests are sparingly used. Suggestions have come from critics of the system that, for a variety of reasons, it would be desirable to make better use of objective tests in our universities. While we consider that the essay type is perhaps more suitable than other types for "testing the ability of the student to have an integrated and connected view of some parts of the subject and also whether the candidate is able to exhibit his knowledge in a constructive and intelligible form", there are other aims the attainment of which are better evaluated by multiple-choice questions, open-book examinations, short answer examinations, *viva-voce* tests etc. In other words the methods of evaluation to be employed should be determined by the objectives to be evaluated, though for some time to come the essay type examination may have to play a prominent role in our examination system, handicapped as our universities are by limitations of various kinds which make any immediate large-scale adoption of other methods impracticable. We have, however, to make sure that the essay type examination is turned into a fitter instrument of assessment than is the case at present in most universities by taking steps to reduce the subjective element and

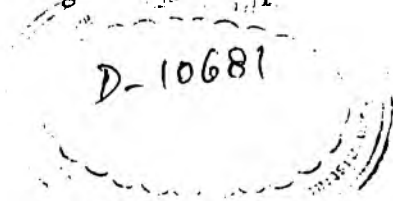
bringing about a greater conformity between the examination and the objectives of teaching different subjects at different levels.

Our zeal for examination reform should not, therefore, lead us to an outright condemnation of the prevalent essay type of examinations. There is a danger that the high standards which we are anxious to maintain or create may be jeopardised by excessive reliance on objective tests. It is noteworthy, in this connection, that in the United States where such tests have been rather widely used, there is considerable concern over the vicious influence that objective testing has exercised over the selection of curricular content, the methods of teaching and the study habits of students. Both the students and the teachers have been not seldom found to respond to the objective test by developing "test wisdom". This includes the use of "guess-papers", cheap made-easy notes and confining oneself to a reading of the texts from which the test items are expected to be culled. It is also true that American technology has tended to transform the time-honoured adventure of examination into yet another technology. We have neither the resources nor technical man-power to do the same. Indeed, no other country in the world has thought it desirable to give up the essay type of examination in favour of large scale objective and standardised tests.

It is desirable to have two examiners to mark each paper at the external examination and to assign to the candidates the average of the marks assigned by them as the final mark. Such a practice would act as a check on the subjective element in the marking of papers by a single examiner. But if it is felt that this procedure would involve excessive expenditure and delay in declaring the results, it may be adopted only at examinations for Master's and Honours degrees and the final professional examinations.

Classification of students

Since the subjective element cannot be completely eliminated from external examinations, it is not desirable within the same class or division to differentiate students according to marks. For the purpose of determining the award of prizes or scholarships the



giving of numerical marks may be useful, and even necessary. But these marks should not be published or made known to candidates, and the published results should only indicate the class or division in which they are placed, without mentioning any order of merit in the same class or division.

The *raison d'être* of classifying successful students in the conventional three classes has been questioned by many competent authorities. For instance, the U.G.C. recommended that the universities should not award a third class at the Master's examination and that at this examination there may be only two classes, the first and the second, which may be awarded on the basis of an aggregate of 60 per cent and 45 per cent of the marks respectively. We are in general agreement with this view, as a third class at the Master's level would appear to be a contradiction in terms. It is, however, a necessary corollary to this reform that standards of the present first and second classes are maintained.

Medium of examinations

Many universities may be changing over from one medium to two or three other media of instruction during the coming years. Certain universities have already permitted their students to answer questions in either English or Hindi or the regional language, and more and more universities are likely to do so in the future. Examinations will encounter a new difficulty when more than one medium is permitted. This arises from the fact that papers written in different languages are likely to be examined by different standards. The following observations made by a committee of the University Grants Commission with regard to the development of certain multi-lingual universities deserve consideration :

“There should be some co-ordinating agency in ensuring that the same standards are maintained in the examinations conducted by the university if there is an option in the media of instruction and examination; if all that is done towards maintaining uniformity of standards in a multi-lingual university is to get a reader well-versed in the language concerned to read out and

translate a script to the examiner, there are bound to be as many standards as there are media of examination recognised by it.”

Particular care will have to be taken and suitable methods devised for co-ordinating evaluations undertaken by examiners of different scripts. The examiners in the different media (at least head examiners) should meet and define the standards and spread of marks in respect of them, before actual evaluation work is undertaken by them. The procedure recommended by the Educational Testing Service, Princeton, U.S.A. in evaluating essay answers may be found useful in this connection (see Appendix XIV). Though the procedure relates to assessing essays written only in one medium, it can be applied to scripts in all the media of examinations of a university.

Administrative problems

Certainly one area in which improvement is urgently called for concerns the administration of examinations. All universities have an Examination Department whose head is either a ‘Controller of Examinations’ or a Deputy or Assistant Registrar. It is not necessary to point out that this department is entrusted with one of the most important functions in the university. The department has to operate with speed, efficiency and secrecy. It is doubtful whether the examination departments in many universities can claim to possess all these virtues. The administration of examinations seems to involve an enormous wastage of time, resulting in late publication of results and consequent difficulties students have to face in regard to obtaining admission to courses in other universities in India or abroad. Sometimes a whole year is lost because of this. The matter needs serious and urgent attention on the part of universities. Steps have to be taken to make the arrangements in the examination departments conducive to dealing with the work in an efficient manner.

It may also be necessary to introduce more modern methods of tabulation and classification of marks than are at present in vogue in a large number of universities where the services of ‘teacher tabulators’ are used for the purpose. It seems desirable to mechanise

at least part of this work in the interest of both speed and secrecy. For secrecy again, fictitious numbers may be given to the answer books before they are despatched to the examiners and papers from different areas mixed up, so that an examiner does not know where the scripts come from nor receive papers from the same area or institution.

Speed also requires that the various processes associated with examinations such as appointment of examiners, allocation of time for setting question papers and for valuation should be well organised. Instances have come to our notice of universities asking examiners to set question papers in less than a week or evaluate a large number of papers in an equally short period. These, in spite of the best efforts of the teachers concerned, are likely to lead to shoddy work on their part.

Summary and Resulting Recommendations

The most compelling reason for examination reform comes from academic considerations. Teaching as well as learning are bound to be adversely affected by a defective examination system since both are closely related to examinations. Examination reform has also assumed a new urgency in the present context of the social and economic life of our country which calls for a better kind of education and a more efficient system of evaluation.

In any serious discussion of examination reform it is necessary to draw a distinction between the educational aspects of the examination system and the technical processes involved in the conduct of examinations. The educational aspects deserve consideration on the ground that examination is an organic part of both teaching and learning. It is not possible to reform any one of these processes in isolation. Thus examination reform entails not only improvement of the mechanics of evaluation but also betterment of the conditions in which teaching and learning are carried on in the universities and colleges.

Educational Issues

One of the most important causes of the wastage of 50 to 60 per cent in individual examinations in India is that universities admit many students who are ill suited for higher education. It is, therefore, necessary to view the problem of admission to colleges and universities objectively and think of it primarily in terms of academic and national needs and in relation to the availability of national resources. The present method of selecting students for

admission does not operate to any great extent as a true selection for university studies, especially in those universities which affiliate a large number of colleges.

Our recommendation with regard to a selective process of admission in the universities and colleges does not however imply that each student is to be chosen for a university course on the results of a battery of aptitude tests. We do, however, emphasise that only those candidates should be admitted to colleges who are capable of profiting by higher education. Admissions may be based either on a cumulative record of the students' performance at school (if such a record is available) or by testing their mastery of the language to be used at the university and of their intellectual maturity and general interests by introducing in the school leaving examination one or two additional papers for this purpose for those who wish to go to universities. Students who do well in this part of the test and at the same time secure qualifying marks in the general examination may prove to be more suitable for admission to the university.

It is necessary to put the examination in its proper perspective as something which contributes to the total purpose of education. It must not be regarded as an end in itself and allowed to dominate the whole educational system. It is, therefore, necessary to re-state the educational objectives underlying examinations. This may be done by groups of teachers in seminars or workshops. When the objectives are defined it is possible to formulate more precisely the specific nature of the knowledge and competence required to be demonstrated in an examination. Such clarification of objectives will help to draw up syllabuses which while indicating the range of material to be studied do not merely enumerate the topics to be studied. A syllabus properly devised should state not only the minimum requirements in a subject but also indicate the wider possibilities of study and learning in the field. It is also necessary continually to revise the syllabi in accordance with the changes that take place in the field of knowledge concerned. This cannot be too strongly stressed.

Reform in examinations cannot be brought about without

improvement in the methods of teaching and learning. If the quality of teaching is to be improved, more time must be devoted to tutorials and seminars, bringing the students more intimately in contact with the mind of the teacher. If this is done the number of lectures could in many cases be reduced with advantage. The value of the tutorial system as an aid to good education in the universities has been widely recognised and its importance in India has also been emphasised. Admittedly, it will be difficult to introduce in our universities tutorials of the kind that prevails in Oxford and Cambridge, but an experiment could be made along the lines of the system adopted in the provincial universities in the U.K. or in the colleges in the U.S.A. where 10 to 15 students are placed in a tutorial group. It should be possible to make a better use of the existing staff and buildings by cutting down the number of lectures by about 50 per cent and introducing tutorial instruction to take its place. The present practice of "covering" the syllabus in lectures could be given up and lectures related to salient aspects of a subject rather than to its details. Together with a system of tutorials, it should also be possible to have larger lecture classes, thus obviating the need for any large-scale expansion of teaching staff.

It is necessary to point out that examinations should periodically reveal to the student his own progress in studies. For effective teaching also it is necessary that the teacher should know at regular intervals how the students are progressing. If the performance of a student is to be judged by a single final examination, the whole purpose of the educative process will tend to be narrowed down by this single end. It is, therefore, necessary to evaluate in some way the advance made by the student periodically. The significance of a system of internal assessment based on a record of class-work can be easily seen, in this context. In the initial stages only a small weightage may be given to internal assessment. If the system is wisely developed weightage could be increased subsequently. While we agree that some difficulties and dangers will have to be met in introducing internal assessment in our universities, we are of the view that the experiment is worth trying and that the educational merits of such assessment outweigh on the whole the risks involved.

An allied consideration relates to the question of having more than one external examination for a degree course. When the course of study involves major and minor subjects, it should be possible for the student to take examination in the minor subjects at an earlier stage, the final examination being limited to the major subject. This arrangement has the advantage of reducing the strain of the examination system on the student and affording him greater scope for intelligent study and preparation. In a number of universities this is already being done.

Technical Questions

One of the most difficult problems associated with the present system of examination is the subjectivity involved in the marking of scripts. The current methods of marking examination scripts and of combining and tabulating marks in university examinations without reference to recognised statistical procedures are not satisfactory. Sometimes moderators are appointed to adjust the difference between the standards of the various examiners but the methods they often use do not have the effect of bringing the marks to a common scale. If it is not possible to use the appropriate statistical methods, it will be necessary at least to define clearly the spread of marks, so that the evaluation may be undertaken on a similar basis by different examiners.

The problem of scaling arises (i) when the examination scripts are so numerous that they have to be divided amongst many examiners and (ii) when the marks in different papers and subjects have to be combined into a final result. Steps will have to be taken to ensure proper scaling in both the situations.

While we agree that the essay type questions are perhaps more superior to other types for testing the ability of the student to have an integrated and connected view of parts of the subject and also where the candidate is able to exhibit his knowledge in a constructive and intelligible form, there are other objectives the attainment of which is better evaluated by true-false and multiple-choice questions, open-book examinations, short answer examinations, *viva-voce* tests etc. Immediate improvement in the examination system, however, seems to lie in the direction of

making the essay type examination less subjective, so that a greater degree of objectivity in evaluation could be secured. One of the ways in which this could be done is to ensure that the examinations are based on well-defined objectives of teaching in the different subjects. It is also desirable to appoint two examiners to mark each paper at the external examination and to assign to the candidates the average of the marks assigned by them as the final marks.

The Committee considers that it is not desirable to grade students marks-wise in view of the difficulty of attaining complete objectivity in regard to evaluation of scripts at present. We also agree with the U.G.C.'s view that the university should not award a third class at the Master's Degree examination and that at this examination there may be only two classes, the first and the second, provided the original standards of the first and second classes are maintained.

Examinations will encounter a new difficulty when more than one medium is permitted. This arises from the fact that examinations in different media bring to bear upon the evaluation of papers different standards. Suitable methods will have to be devised for co-ordinating the evaluation undertaken by examiners of different scripts. The examiners in the different media should meet and define the standards and spread of marks in respect of them before actual evaluation work is taken up.

Improvement is urgently called for in respect of the administration of examinations. There appears to be at present an enormous wastage of time in the administrative work pertaining to examinations which is responsible for late publication of results etc. It will be desirable to modernise methods of tabulation and classification of marks in the interest of speed and accuracy. Speed also requires that the various steps to be taken for the conduct of examinations, such as appointment of examiners, allocation of time for setting question papers and for evaluation should be well organised.

Specific Recommendations

1. No reform in the system of examinations will reduce the failure rate in our universities and colleges, unless the prevailing

admission procedures are improved. We have therefore to see that only those candidates are admitted to universities who can profit by higher education. One of the feasible ways by which this could be done is to introduce in the School Leaving Examination two additional papers, one to test competence in the use of the language of the university and one to test intellectual maturity, for those who wish to enter the university.

2. Teaching work should be done not only through lectures but through tutorials, seminars, etc. It will be desirable to hold periodical short tests on the work done in the tutorials and to maintain a record of the assessments made. This should be regularly evaluated. Each university may decide what weightage should be given to this. In order to make room for more tutorials, lectures may be cut down (it should be possible to reduce them by 50 per cent) and the teaching work divided between tutorials and lectures.

3. The U.G.C. should encourage seminars, discussions and conferences of university and college teachers for defining the objectives of teaching and examinations in different subjects at various levels. A clear conception of the aims of teaching will facilitate good teaching and bring about a greater conformity between examinations and teaching.

4. Research should be undertaken in regard to both the educational and technical aspects of examinations. Topics which may be taken up for research in this connection are indicated in the report. It should be possible for the departments of education in universities to undertake such work as a part of their normal activities. Perhaps the newly created National Council for Educational Research & Training would also be able to assist in this. There should be arrangement in the University Grants Commission to co-ordinate the research work of the different universities and to disseminate information and conclusions with regard to the studies undertaken in the country.

5. Where the courses are spread over two years or more, convenient spacing of the examinations should be arranged, so that the examination is not concentrated at the end of the final year.

6. Use of methods of evaluation other than essay type examination such as multiple-choice tests, short answer tests, open-book tests, *viva-voce*, etc. may be tried wherever necessary. While, for a variety of reasons, the essay type examination may continue as the chief mode of evaluation in our universities, it is necessary to make it a fitter instrument for measuring the educational development of the students.

7. The present methods of marking examination scripts and of combining and tabulating marks in university examinations without reference to recognised statistical procedures are not satisfactory. The procedures will have to be developed to make marking and combining of marks more objective. Suggestions in this regard are given in the report.

8. In view of the difficulty of achieving objectivity and precision in the marking of papers, it may be desirable to rank students class-wise rather than marks-wise. In general, only two classes may be awarded for the Master's Degree, *viz.* the first and the second, provided the standards of the present first and second classes are maintained.

9. In universities where more than one media of examination exists, the examiners in the different media (at least head examiners) should meet and define the standards and spread of marks to be adopted in the evaluation work undertaken by them, in order to avoid variation in the marking of scripts.

10. Ways and means must be found for avoiding wastage of time in the administrative work of examinations resulting in late publication of results and consequent difficulties to students.

Sd. S. R. Dongkerkery (*Chairman*)
 Sd. D. C. Pavate
 Sd. H. J. Taylor
 Sd. Salamat Ullah
 Sd. T. K. N. Menon
 Sd. E. A. Pires
 Sd. S. Mathai
 Sd. P. J. Philip (*Member-Secretary*)

APPENDIX I

Comments of the Vice-Chancellor's Conference on the First Draft Report of the University Grants Commission Committee on Examination Reform

(1) The Committee thought that it was necessary to recognize examinations as an important element in the teaching process, contributing to the whole purpose of education and, in this context, it was necessary to define the educational objectives of teaching different subjects in addition to the general objectives of a liberal education. These objectives should determine the teaching as well as the examination in various subjects.

(2) That there is need for a continual investigation of the operation of the existing system of examinations in Indian Universities, particularly with regard to the following:

- (a) Reliability and validity of the present examinations;
- (b) Techniques of evaluating, scaling and combining of marks;
- (c) Methods of selection of students for university courses and development of aptitude tests;
- (d) Questions set in different subjects in examinations;
- (e) Development of material for objective testing in various subjects fields;
- (f) Survey of the methods available for the internal assessment of students.

(3) The Committee was generally of the opinion that the important thing to be achieved in our universities and colleges is that students do regular and continuous work throughout the course,

and that for this purpose detailed cumulative records should be maintained by the College showing attendance at lectures, tutorials, discussions and libraries and the carrying out of other forms of prescribed work. In addition to this a record should be maintained of the evaluation of periodic tests conducted by the teachers in each subject.

(4) The total marks based on such records, which may be reckoned at about 20 per cent of the total marks in a subject or paper, should be taken into account in determining the final result, including the classification of students. In order to secure regularity and objectivity in the maintenance of these records the Committee thought it would be desirable to insist on such records being sent by the colleges to the universities at the end of every term. The maintenance of these records by the colleges should be constantly kept under examination by the universities. In case of University Departments, the record of term work should be sent by the teacher concerned or the head of the Department to the Registrar, Controller of Examinations or other officer appointed for this purpose.

(5) The Committee thought that the maintenance of cumulative records which involved a large number of teachers, and the communication of these records to the University will largely guard against abuses. In addition to this, it may be necessary for the University to examine cases of wide discrepancies between the marks given by the Colleges for term work and the marks obtained by the students in the University Examinations. The maintenance of cumulative records should be undertaken in the first instance by central and unitary universities but any request from teaching and affiliating universities for financial assistance for trying out this experiment should be treated on the same basis as for the central and unitary universities.

(6) In the opinion of the Committee it is essential that the syllabuses in the various subjects should be revised from time to time so as to include therein the results of new developments, and also in the light of the fresh experience and the redefinition of the educational objectives.

(7) It is important that the work of the teachers should not be confined to lectures, but should include tutorials, discussions, seminars, etc. and that attendance at all these should be considered equally important and recorded.

(8) Each university should set up a special unit to study examination reform in the context of that university. The unit should include at least one expert with a knowledge of modern evaluation techniques and statistical methods. The unit would study the design of examinations, the tabulation and scaling of marks, and other features of the examination procedure, and make recommendations to the University authorities. Senior teachers and those with special competence, should be associated with the unit.

(9) The University Grants Commission should take steps to plan and initiate reforms and to co-ordinate the work done in this field in the various universities. It would also be necessary to provide facilities for rapid statistical work and the production of text material. The information and conclusions of such work should be made available to universities and a journal may be brought out for that purpose.

APPENDIX II

TABLE I

Pass Percentage in I.A., B.A., I.Sc., & B.Sc. Examinations

<i>Year</i>	<i>I.A.</i>		<i>B.A.</i>		<i>I.Sc.</i>		<i>B.Sc.</i>	
	<i>Pass%</i>	<i>Fail%</i>	<i>Pass%</i>	<i>Fail%</i>	<i>Pass%</i>	<i>Fail%</i>	<i>Pass%</i>	<i>Fail%</i>
1951	42.3	57.7	47.5	52.5	43.2	56.8	48.9	51.1
1952	41.5	58.5	51.7	48.3	39.0	61.0	45.9	54.1
1953	43.3	56.7	45.5	54.5	42.5	57.5	44.2	55.8
1954	42.0	58.0	44.9	55.1	41.8	58.2	46.1	53.9
1955	45.7	54.3	48.5	51.5	46.3	53.7	45.9	54.1
1956	44.4	55.6	45.6	54.4	44.3	55.7	47.1	52.9
1957	44.2	55.8	47.9	52.1	42.7	57.3	46.2	53.8
1958	41.3	58.7	47.8	52.2	43.4	56.6	48.6	51.4
1959	N.A.	N.A.	46.7	53.3	N.A.	N.A.	55.6 est.	44.4 est.
1960	N.A.	N.A.	43.0	57.0	N.A.	N.A.	51.0	49.0

TABLE II
Pass Percentage in I.Com. & B.Com. Examinations

<i>Year</i>	<i>I.Com.</i>		<i>B.Com.</i>	
	<i>Pass%</i>	<i>Fail%</i>	<i>Pass%</i>	<i>Fail%</i>
1951	51.7	48.3	47.5	52.5
1952	51.5	48.5	50.5	49.5
1953	48.9	51.1	45.6	54.4
1954	46.5	53.5	45.5	54.5
1955	49.3	50.7	47.2	52.8
1956	46.7	53.3	46.1	53.9
1957	44.2	55.8	49.5	50.5
1958	44.8	55.2	46.0	54.0
1959	N.A.	N.A.	49.1	50.8
1960	N.A.	N.A.	49.2	50.8

APPENDIX III

TABLE I
Percentage of Failures in M.A. Examination

<i>Year</i>	<i>Number appeared</i>	<i>Number passed</i>	<i>Pass Percentage</i>	<i>Percentage of failures</i>
1948-49	4,427	3,343	75.5	24.5
1949-50	5,992	4,503	75.2	24.8
1950-51	8,123	5,969	73.5	26.5
1951-52	8,404	6,467	77.0	23.0
1952-53	9,256	7,038	76.0	24.0
1953-54	10,488	7,889	75.2	24.8
1954-55	11,754	8,886	75.6	24.4
1955-56	13,630	9,528	69.9	30.1
1956-57	13,009	10,483	80.6	19.4
1957-58	14,355	11,670	81.3	18.7
1958-59	17,002	13,837	81.4	18.6
1959-60	20,861	16,048	76.9	23.1

TABLE II
Percentage of Failures in M.Sc. Examination

<i>Year</i>	<i>Number appeared</i>	<i>Number passed</i>	<i>Pass percentage</i>	<i>Percentage of failures</i>
1948-49	1,121	828	73.9	26.1
1949-50	1,440	1,078	74.9	25.1
1950-51	1,723	1,398	81.1	18.9
1951-52	2,085	1,641	78.7	21.3
1952-53	2,234	1,780	79.7	20.3
1953-54	2,772	2,146	77.4	22.6
1954-55	3,108	2,348	75.6	24.4
1955-56	3,262	2,519	77.2	22.8
1956-57	3,646	2,928	80.3	19.7
1957-58	3,723	2,942	79.0	21.0
1958-59	4,214	3,400	80.7	19.3
1959-60	4,672	3,672	78.6	21.4

TABLE III
Analysis of Examination Results

<i>Percentage of</i>	<i>M.A. examination</i>		<i>M.Sc. examination</i>	
	1958-59	1959-60	1958-59	1959-60
I Classes	5.0	4.6	28.0	27.3
II Classes	40.3	39.9	53.9	57.4
III Classes	54.7	55.5	18.1	15.3

APPENDIX IV

Questionnaire of the University Grants Commission

No. F. 60-1/59(H)

Dated the January, 1960
Paus, 1881

From

The Secretary,
University Grants Commission.

To

The Registrar,
.....
.....

Subject: *Reform of the examination system in
Indian Universities.*

Sir,

I am directed to request you to furnish information on the following points to the University Grants Commission urgently in connection with the question of reform of the examination system:

I. *Nature of Examination:*

- (a) Are all examinations conducted by the University written, apart from 'practicals' in Science?
- (b) What are the other methods adopted, if any?

II. *Internal Assessment:*

- (a) Are any (credits) given for the work of students on the basis of class tests, writing of essays, practicals etc.?

- (b) Are such 'credits' taken into account in the final examination and, if so, to what extent (percentage)?

III. *Question Papers:*

- (a) Are question papers usually set by one examiner or more than one examiner?
- (b) Are question papers 'moderated' and, if so, why is it considered necessary?
- (c) What is the nature of the machinery set up for moderation of question papers?

IV. *Marking Papers:*

- (a) Are papers marked by one examiner or by more than one examiner? If the marking differ, when there are two or more examiners, how is the final mark arrived at?
- (b) After the marks have been submitted by the different examiners, are they scaled by some suitable procedure in order to bring them to a common standard?
- (c) Are examiners given any instruction regarding spread of marks, such as, 'very good answers' may be given 75 per cent and upward, 'good answers' may form 50 per cent to 70 per cent 'average answers' from 30 per cent to 50 per cent and 'bad answers' below 30 per cent?
- (d) Are students expected to obtain minimum pass marks in the individual papers of the same subject, if there are more than one Paper in a subject?
- (e) How many answer books are generally given to a single examiner for valuation?

V. *Classification of Results:*

- (a) In what classes are the successful candidates placed?
- (b) If the classification is into first, second and third, how are they defined in terms of marks?

- (c) If the successful candidates are classified differently, how is this done?

VI. *Research in Examination :*

- (a) Is any research work being done in your University on the examination system?
- (b) If so, what are the problems on which research work is undertaken (example of research with publications, if any, may be sent);
- (c) Is there any 'evaluation or psychometric unit' in your University for providing necessary assistance in matters relating to examinations and for promoting research in this field? If so, when was it started and what is the nature of its work?

VII. *Reform in Examination System :*

- (a) Has the University introduced any reform in the examination system?
- (b) If so, what reforms have been effected and how have they influenced the operation of the examination system;
- (c) Has there been any seminars or organised discussions on problems relating to the examination system in your University?

The University Grants Commission would make a special request to your University kindly to give the highest priority to sending answers to these questions. The answers may be brief. The Commission earnestly hopes that necessary information will be supplied by your University as early as possible, in any case not later than the 7th February, 1960.

Thanking you,

Yours faithfully,

(P. J. Philip)
for Secretary

APPENDIX V

Tabular Presentation of the Questionnaire Study

Dr. A. K. Gayan

Indian Institute of Technology, Kharagpur

Note: Please see the following charts for Appendix V.

Table I: Summary of the Nature of Examinations obtaining in Different Universities

[Based on data supplied in reply to Q. 1 (b)]

Sl. No.	Name of University.	All Examinations written or not [Q. 1 (a)]	Other Examination Methods							Other Exams. Medicine, Vet. Sc.	
			Under-graduate			Post-graduate					
			Arts	Sc. Engg. Tech.	Edn. Com. Hns. Sc. Phy. Edn.	Arts, fine Arts.	M.Sc. Ph.D.	Socio, Law, Language, Edn.	Ag. Com.		
1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	
1.	Agra	Y					T	T	F	T	R,T,O
2.	Aligarh	Y									
3.	Allahabad	Y					O	O			
4.	Andhra	Y					T,O	T,O	E		C,O
5.	Annamalai	Y					T	T			
6.	Banaras	Y		S,O				S,O			
7.	Baroda	Y		S			O	O,T	F,T,O	O,E	T,O
8.	Bombay	Y	O	S,O			T,O	T,O	T	T,O,P	O
9.	Calcutta	Y									O
10.	Delhi	Y		S,O							O
11.	Gauhati	Y		O			T	T,O			
12.	Gorakhpur	Y				P	O	O	T,O		
13.	Gujarat	Y									
14.	Jadavpur	Y		O				O			
15.	Jammu & Kashmir	Y									
16.	Karnatak	Y	S	S,O				O			
17.	Kerala	Y					O				
18.	Kuruksetra	Y									
19.	Lucknow	Y	P	O			P,O			O	
20.	Madras	Y		S	S		S	S,O			C,O
21.	Marathwada	Y									
22.	Nagpur	Y									
23.	Osmania	Y									
24.	Panjab	Y				O,P	T,P		T,P	T,O	O T,P
25.	Patna	Y									
26.	Poona	Y									
27.	Rajasthan	Y									
28.	Roorkee	Y									
29.	S.V. Vidyapeeth	Y									
30.	Saugar	Y	P			O	P,O			O	
31.	S. N. D. T. (W)	Y									
32.	Sri Venkateswara	Y					O	O			
33.	Utkal	Y									
34.	Visva Bharati	Y		S	O		O,S,P				
35.	Varanaseya Sanskrit	Y			P		O,P				
36.	Vikram	Y					O	O			
37.	I.I.S.	Y									
38.	I.A.R.I.	N									

Explanation of Symbols:

(Y) stands for Yes and (N) for No; (O) stands for Oral or *Viva-voce* examinations, (T) for Thesis or Dissertation; (P) for Practical in other than Science subjects; (S) for Sessional work or Term work or Class work; (F) for Field work or Project work; (E) for Special Essays or Reports; (C) for Clinical and (R) for Research work.

Table II: Summary of Information about the Credit given to Internal Assessment in the Final Examination of Different Universities

[Based on data supplied in reply to Q. II (a) and (b)]

Sl. No.	Name of University.	I.Sc.	B.Sc.	B.A.	M.Sc.	M.A.	B.Com.	B.Ed.	M.Ed.	LL.B.	B.E. Engg. & Tech.	M.B.B.S. Vet. & B. Phar.	T.D.	Dip. in Lib.	Dip. Craft	Pre- Univ.	Pre- Profess.	Prac. Exam. of Var. Course	Explanation of Symbols.
1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.
1.	Agra		10/20								20								
2.	Aligarh		5	20				35	25		30				37				(a) 15% for College Test 5% for Attendance.
3.	Allahabad		10		10														
4.	Andhra		20					50			17/30	29							
5.	Annamalai		8		14						23						23		(b) For Class Tests and Practicals.
6.	Banaras										37/43	40							
7.	Baroda		30	30			30	30			30	30							
8.	Bombay							20			30	30/40							(c) For Sessional in Science only.
9.	Calcutta	10	10		10														
10.	Delhi	20	20		20			25	25		20/40	20						20/50	
11.	Gauhati	10	10		10						10/100	10							(d) Practice already there to be abolished.
12.	Gorakhpur		10					25											
13.	Gujarat		20											30/40					
14.	Jadavpur		—								30								(e) Passing of Tutorials and Term work essential. No figure quoted.
15.	Jammu & Kashmir		20	20	20	20	20	20			20(a)								
16.	Karnatak																		
17.	Kerala		20		20						(f)	45				20			
18.	Kuruksetra				(.....f.....)														(f) Candidates are to take Viva in all examinations including Practical.
19.	Lucknow		20/25	20/25			20/25			20/25		50							
20.	Madras	(Credit given in certain subjects only to a small extent—proportion not mentioned)																	
21.	Marathwada		(.....20b)																
22.	Nagpur																		
23.	Osmania		20	20			20									20			(g) Whole of Sessional Credited in Final Exams.
24.	Panjab															20/25			
25.	Patna		10		20														
26.	Poona																		
27.	Rajasthan										7	10							(h) In Kerala for Practical Records 3 to 7% and for Sessionals about 40% in Engg. Courses, but about 15% in B.A./B.Sc. Courses.
28.	Roorkee		(.....g.....)																
29.	S.V. Vidyapeeth		(.....e.....)																
30.	Saugar	20	20		20(e)														
31.	S. N. D. T. (W)							40											
32.	Sri Venkateswara					5/17													
33.	Utkal								5		100	5				20		5	
34.	Visva Bharati			20		16		25							25				
35.	Varanaseya Sanskrit		(.....d.....)																
36.	Vikram																		
37.	I.I.S.		(.....50.....)																
38.	I.A.R.I.	(Credit given but proportion not mentioned)																	

Table III: Summary of Information about the Paper-setting and the Moderation work of the Different Universities

[Based on data supplied in reply to Q. III (a) to III (c)]

Sl. No.	Name of University.	III(a)				III(b)	III(c)	Explanation of Symbols
		No. of Paper Setter				Whether Moderated	Moderation Machinery	
		U.G. Edn.	P.G.	M.B.B.S.	B. Eng.			
1.	Agra	1	1	2	2	Y	B4	(i) B or Board stands for a 'Board of paper-setters' or 'Two or more paper-setters' set the paper(s) jointly.
2.	Aligarh	1	1	1	1	Y	B3	
3.	Allahabad	1	1	1	1	Y	B	
4.	Andhra	B	1	B	B	Y	B3	(ii) Y(N) for 'Yes when Necessary', Y(UG) for 'Yes in U.G.' and Y(BE) for 'Yes in Engineering Subjects only'.
5.	Annamalāi		(BOARD)			N	—	
6.	Banaras	1	1	1	1	Y	B3	
7.	Baroda		(BOARD)			N	—	(ii) Y(N) for 'Yes when Necessary', Y(UG) for 'Yes in U.G.' and Y(BE) for 'Yes in Engineering Subjects only'.
8.	Bombay		(BOARD)			N	—	
9.	Calcutta	1	2			Y	2	
10.	Delhi	1	1	1	1	Y	B3	(iii) B, B2, B3, B4 mean Moderation Machinery consists of a Board, of 2, 3, 4..... members with or without the Head of Department as Chairman.
11.	Gauhati	1	1	1	1	Y	2	
12.	Gorakhpur	1	1	1	1	Y	B4	
13.	Gujarat		(BOARD)			N	—	(iii) B, B2, B3, B4 mean Moderation Machinery consists of a Board, of 2, 3, 4..... members with or without the Head of Department as Chairman.
14.	Jadavpur	1	1	1	1	Y(N)	B	
15.	Jammu & Kashmir	1	1	1	1	N	—	
16.	Karnatak		(BOARD)			Y(N)	—	(iii) B, B2, B3, B4 mean Moderation Machinery consists of a Board, of 2, 3, 4..... members with or without the Head of Department as Chairman.
17.	Kerala		(BOARD)			Y	B2	
18.	Kuruksetra	1	1	1	1	N	—	
19.	Lucknow	1	1	2	1	Y	B3	(iii) B, B2, B3, B4 mean Moderation Machinery consists of a Board, of 2, 3, 4..... members with or without the Head of Department as Chairman.
20.	Madras	1	1	1	1	Y	B	
21.	Marathwada		(BOARD)			N	—	
22.	Nagpur	1	1	1	1	Y	B3	(iii) B, B2, B3, B4 mean Moderation Machinery consists of a Board, of 2, 3, 4..... members with or without the Head of Department as Chairman.
23.	Osmania	1	1	2	2	Y	B4	
24.	Panjab	1	1	1	1	N	—	
25.	Patna	1	1	2	1	Y	B4	(iii) B, B2, B3, B4 mean Moderation Machinery consists of a Board, of 2, 3, 4..... members with or without the Head of Department as Chairman.
26.	Poona		(BOARD)			N	—	
27.	Rajasthan	1	1	1	1	Y(BE)	B3	
28.	Roorkee	1	2			Y(UG)	B4	(iii) B, B2, B3, B4 mean Moderation Machinery consists of a Board, of 2, 3, 4..... members with or without the Head of Department as Chairman.
29.	S.V. Vidyapeeth	1	1	1	1	N	—	
30.	Saugar	1	1	1	1	Y	B3	
31.	S. N. D. T. (W)		(BOARD)			N	—	(iii) B, B2, B3, B4 mean Moderation Machinery consists of a Board, of 2, 3, 4..... members with or without the Head of Department as Chairman.
32.	Sri Venkateswara		(BOARD)			Y(N)	—	
33.	Utkal	1	1	1	1	Y	B3	
34.	Visva Bharati	1	2	1	1	Y	B2	(iii) B, B2, B3, B4 mean Moderation Machinery consists of a Board, of 2, 3, 4..... members with or without the Head of Department as Chairman.
35.	Varanaseya Sanskrit	1	1	1	1	Y	B	
36.	Vikram	1	1	B	B	Y	B2	
37.	I.I.S.	1	1	1	1	N	—	(iii) B, B2, B3, B4 mean Moderation Machinery consists of a Board, of 2, 3, 4..... members with or without the Head of Department as Chairman.
38.	I.A.R.I.		(BOARD)			N	—	

Table IV: Summary of Information about MARKING OF PAPERS in the Different Examinations of Universities

[Based on data supplied in reply to Q. IV (a) to IV (e)]

Sl. No.	Name of University.	No. of Examiners Q. IV(a)			If scaled common standard Q. IV(b)	Instrn. Mark Spreading Q. IV(c)	All Exams.	Passing Individual Papers or subjects as a whole Q. IV (d)				No. of scripts for an Examiner Q. IV(e)	Explanation of Symbols.
		P.G. or Hons. Exams.	U. G. & other Exams.	Marks com- bining Method.				Under Graduate	P.G./ Law/ Edn.	Science: Theory & Practical.	Engg./ Tech./ Med.		
1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.
1.	Agra	1	1	—	N	Yg(s)	N°					400(M)	For Question IV(a) in Cols. (3), (4) & (5): 2° stands for 2 examiners in cases of M.Ed. thesis only; 1°, for more than 2 examiners only in M.B.B.S. Examination; X, for the case when the question is not properly understood; M for mean of Awards as Final Mark on the paper when there is a difference below 10%, M† for Mean of 3 examiners when difference is above 12%; M†† for Mean of 3 examiners when difference is above 15%; M††† for Mean of 3 examiners when difference is above 20%. For Question IV(b) in Col. (6): Y for Yes, but with meaning of 'scaling' quite different from what the question actually implies. N, for No scaling for common standard done, (?) for Mis-understanding of the question. For question IV(c) in Col. 7: Y stands for instruction without any specific mark partitioning to describe different qualities of performances; Yg for instructions on grading of students in different classes (divisions) on the whole examination; Yc for instructions on marking of scripts only: Yg(s) for instructions on grading but with mark partitioning different from those of the question-iten; Y° for instruction on U.G. Examination by Head Examiner; Yg(w) for instructions on grading without numerical marking; Ym for instruction only when there is more than one examiner; N for 'no instruction' and Ny for 'little instruction'. For question IV(d) Cols. (8), (9), (10), (11), (12) & (13): Y stands for Yes; Y° for Yes but not in all cases; Y°° for Yes in special paper only; Y? for yes but if one gets second class in aggregate then he is passed and that even if he fails in a part paper; N, for No; N°, for No except in few examinations (not mentioned). For Question IV (e) Col. (13): (O) stands for 'ordinarily' or generally, (M) for Maximum; and h for 'Hours'.
2.	Aligarh	X	X	X	N	NY		N	Y			250(M)	
3.	Allahabad	2°	1	M	N	N	N					300(M)	
4.	Andhra	2	1	M	N	Y	N					400(O)	
5.	Annamalai	2	1	M	N	N					Y	500(M-UG)	
6.	Banaras	1	1	—	?	Ny	N				Y	500(M)	
7.	Baroda	2	1	M	N	Ny Yg	N	Y	Y			200(O)	
8.	Bombay	1	1	—	N	N	N		Y				
9.	Calcutta	2/3	1	M	N	Y°	N			Y		500(M-UG)	
10.	Delhi	1	1	—	N	N		N	Y			300(N)	
11.	Gauhati	1	1	—	Y	Y°	N°					300(O)	
12.	Gorakhpur	1	1	—	N	Yg	N					200(M)	
13.	Gujarat	2	1	M	Y	Ny	N		Y			500(M)	
14.	Jadavpur	2	2	M°	N	N	Y				Y	100(M)	
15.	Jammu & Kashmir	X	X	X	Y	N	N			Y		500(M)	
16.	Karnatak	2	1	M	N	N	N					300(O)	
17.	Kerala	2	1	M	Y	Yg(s)	N					400(M-3h) 500(M-2h) 30(M)	
18.	Kuruksetra	1	1	—	Y	N	Y					300(O)	
19.	Lucknow	2°	1	M	N	N	N			Y		500(M)	
20.	Madras	2	1	M†	N	N	Y°					500(M)	
21.	Marathwada	1	1	—	N	N	N					500(M)	
22.	Nagpur	1	1	—	N	Ym	N			Y		300(M)	
23.	Osmania	2	1	M	Y	N	N			Y		300(O)	
24.	Pan'ab	2	1	M	Y	Yg(s)	Y		Y?			350(O)	
25.	Patna	1	1	—	Y	Y	N				Y	200(M)	
26.	Poona	1	1	—	Y	N	Y		Y			500 1000 Sec.	
27.	Rajasthan	2°	1	M††	Y	Y	N					300(M)	
28.	Roorkee	2	1	M†††	N	N	Y					200(M)	
29.	S.V. Vidyapeeth	2	1	M	N	N	N	Y				225(O)	
30.	Saugar	X	X	X	Y	Yg(s)	N		Y		Y	350(M)	
31.	S. N. D. T. (W)	X	X	X	N	N	N	Y				400(M)	
32.	Sri Venkateswara	2	1	M	N	N	N					400(M)	
33.	Utkal	1°	1	M	Y	Y	N					200(M)	
34.	Visva Bharati	2°	1	M	N	N	Y					1 to 80	
35.	Varanaseya Sanskrit	1	1	—	Y	N	Y					500(M), 400(O)	
36.	Vikram	1	1	—	N	Y	N				Y	300(M)	
37.	I.I.S.	1	1	—	N	N	N					10 to 60	
38.	I.A.R.I.	1	1	—	Y	Yg(w)	N					Not mentioned	

Table VI: Summary of Information about the Research in Examination of Different Universities

[Based on data supplied in reply to Q. VI (a) to Q. VI (c)]

Sl. No.	Name of University.	Any Res. work Q. VI(a) (Yes/No)	Res. on what Q. VI(b)	Psycho Metric Unit? Q. VI(c)	Explanation of Symbols.
1.	Agra	N	N	N	Y—Yes.
2.	Aligarh	Y	IES	Y	
3.	Allahabad	N	N	N	N—No.
4.	Andhra	N	N	N	
5.	Annamalai	N	N	N	N°—No but evaluation
6.	Banaras	N	N	N	done from time to
7.	Baroda	Y	RIE	CSS	time.
8.	Bombay	Y	EE	N	
9.	Calcutta	Y	SMRR	N	NM—Not mentioned.
10.	Delhi	N	N	N	
11.	Gauhati	N	N	N	SMRR—Whether sys-
12.	Gorakhpur	N	N	N	tem of marking en-
13.	Gujarat	N	N	N	sures Right Results.
14.	Jadavpur	N	N	N	
15.	Jammu & Kashmir	N	N	N	IES—To suggest ways
16.	Karnatak	N	N	N	and means for im-
17.	Kerala	N	N	N	provement in Exami-
18.	Kuruksetra	N	N	N	nation System.
19.	Lucknow	N	N	N	
20.	Madras	N°	IES	N	EE—Enquiry about
21.	Marathwada	N	N	N	Examination.
22.	Nagpur	N	N	N	
23.	Osmania	N	N	N	TDRIP—Teaching De-
24.	Panjab	N	N	N	monstration and Re-
25.	Patna	N	N	N	search in Industrial
26.	Poona	N	N	N	Psychology.
27.	Rajasthan	N	N	Y	
28.	Roorkee	N	N	N	AP—(PG) — American
29.	S.V. Vidyapeeth	N	N	N	Pattern of Examina-
30.	Saugar	Y	IES	NM	tion in P.G. Courses.
31.	S. N. D. T. (W)	N	N	N	
32.	Sri Venkateswara	N	N	N	CSS — Contemplated
33.	Utkal	N	N	N	Scheme sent.
34.	Visva Bharati	N	N	N	
35.	Varanaseya Sanskrit	N	N	N	RIE—Correlation be-
36.	Vikram	N	N	N	tween scores on In-
37.	I.I.S.	Y	TDRIP	Y	ternal & External
38.	I.A.R.I.	Y	AP (PG)	NM	Examinations.

Table VII: Summary of Information about the Reforms in Examination System of Different Universities

[Based on data supplied in reply to Q. VII (a) to Q. VII (c)]

Sl. No.	Name of University.	Any Reform/ Effect Q. VII(a & b)	Any Seminar or so? Q. VII(c)	Explanation of Symbols.
1.	Agra	N	N	Y—Yes.
2.	Aligarh	Y	SSS	
3.	Allahabad	N	N	Y*—Yes Grade system.
4.	Andhra	Y	N	
5.	Annamalai	N	N	N—No.
6.	Banaras	N	N	
7.	Baroda	Y	SDS	C—Committee
8.	Bombay	Y	D	
9.	Calcutta	N	N	D—Discussion.
10.	Delhi	Y	D	
11.	Gauhati	N	SD	S—Seminar
12.	Gorakhpur	N	N	
13.	Gujarat	N	C	SD—Seminar and Dis- cussion.
14.	Jadavpur	N	N	
15.	Jammu & Kashmir	Y	N	
16.	Karnatak	Y	D	SS—Seminar held twice.
17.	Kerala	Y	N	
18.	Kuruksetra	Y	N	SSS—Seminar held thrice.
19.	Lucknow	Y	N	
20.	Madras	Y*	N	
21.	Marathwada	N	N	SDS—Seminar twice and discussion once.
22.	Nagpur	N	N	
23.	Osmania	N	N	
24.	Panjab	Y	N	
25.	Patna	Y	N	
26.	Poona	N	SS	
27.	Rajasthan	Y	N	
28.	Roorkee	N	N	
29.	S.V. Vidyapeeth	Y	N	
30.	Saugar	Y	N	
31.	S. N. D. T. (W)	N	N	
32.	Sri Venkateswara	Y	N	
33.	Utkal	N	N	
34.	Visva Bharati	Y	N	
35.	Varanaseya Sanskrit	Y	N	
36.	Vikram	N	N	
37.	I.I.S.	Y	D	
38.	I.A.R.I.	N	N	

APPENDIX VI

A Review of Admission Procedures Used in Colleges and Universities Abroad

Dr. Rhea S. Das

Introduction

Technical and scientific manpower requirements and an increasing number of university applicants are among the factors which point towards the need for an evaluation of college and university admission procedures in India. In order to formulate principles and methods for university admissions, a factual analysis of current practices in India would be of value. A survey of admission procedures used abroad might also assist in the formulation of principles and development of methods for future use in colleges and universities. The present review has been undertaken with these objectives in view.

Selection of Placement

Prior to considering procedures used for admission purposes, the nature of the admission decision may be briefly examined. The simplest type of admission decision is one of selection, *i.e.*, an applicant is either rejected or selected. A more complicated type of decision is that of placement, in which an applicant is both accepted and placed into one of several possible courses. The choice of admission procedures to be employed will depend in part upon whether selection or placement decisions are to be made. The relevance of this consideration for university and college admissions in India may be demonstrated by a brief reference to the following trends. Among persons who have attended or completed college and university courses, un-employment is relatively higher among those enrolled in arts and commerce courses than it

is among those who were enrolled in science, technical and vocational courses. Yet, at the present time, student enrolments are increasing more in arts and commerce courses than they are in the latter group of courses. The trends of present and future needs for technical and scientific manpower are also well-known. How the students enter the various types of courses, designated broadly as arts, commerce and science, and whether they enter courses which will most benefit themselves and the country, are questions which are basic to the admission procedures. In India, placement and selection decisions are both of basic importance, however, relatively greater emphasis on placement may be of value over a period of time.

Admission Procedures Abroad

Selection and placement of applicants to institutions of higher learning has been recognised as a problem in many countries. The procedures which they have used and their accumulated experiences with them may be appropriately considered in the context of admission problems in India. For this purpose, attention may be restricted to common admission practices in four industrial countries, viz., Japan, the United Kingdom, the United States of America and the Union of Socialist Soviet Republics.

In Japan, the first step affecting university admission is taken at the end of the compulsory lower secondary school course, when pupils apply for admission to the upper secondary school which is a pre-requisite for college and university education. The pupils who successfully pass achievement tests held for this purpose and who have satisfactory lower secondary school reports are admitted to the upper secondary course. After completion of the upper secondary course, the second step in deciding university admission is taken. It consists of achievement tests conducted by the university and an evaluation of the upper secondary school record, on the basis of which admission decisions are made.

Admission to grammar schools, which generally lead to college and university education in U.K., is limited by the "11+" examination conducted when the pupils are 11 years old. Pupils who are

admitted and complete the grammar school course appear for a school final examination, such as that held for the General Certificate of Education or for the Scottish Senior Leaving Certificate. The subjects passed, and the level of the passes, are taken into consideration by universities in admitting students. Some universities, such as Oxford and Cambridge, and departments within colleges such as University College, London, also hold entrance examinations of the essay type and interviews before admitting students.

Enrolment in the secondary school course preparing pupils for higher education is not limited in the U.S.A. Although different courses are offered in secondary schools, enrolment in them is a matter of individual choice. Successful completion of the secondary school course, in terms of average marks received throughout the course and completion of required subjects, is usually necessary for admission to colleges and universities. Admission procedures depend on the individual institutions of higher learning. State colleges and universities, supported mainly by taxes, are usually required by law to accept all applicants who have graduated (*i.e.*, passed) from an accredited secondary school of that state. For these colleges and universities, entrance examinations are not held for admission purposes but tests are often given entering students for placement advice. Selection and placement of students is often carried out by private colleges and universities, which are supported mainly by tuition fees and endowments. Secondary school record is considered for admission purposes along with scores on aptitude and achievement tests. These tests are often administered on a nation-wide basis by private organisations, an example, of which is the College Entrance Examination Board. Sometimes colleges and universities also conduct interviews for admission purposes.

State examinations are held at the end of Class VII in U.S.S.R. at which time pupils are placed in different secondary schools. Only one of these schools, the 10-year-school, generally prepares the pupils for higher education, although outstanding graduates of other schools are also admitted to institutions of higher learning. Secondary school performance is evaluated in two ways, by marks given on the basis of class performance throughout the secondary school

course, and by marks in the matriculation examination which is held on a state-wide basis. Both school record and matriculation examination are considered for university admission purposes. In addition, the universities hold entrance examinations which most of the applicants are required to take. Matriculation and university entrance examinations consist of both written essays and oral examinations.

On the basis of the overview of procedures used in these four representative industrial countries, it is possible to classify the data used for college and university admissions into the following five types:

- (i) performance on tests prior to entering secondary schools;
- (ii) secondary school performance;
- (iii) performance in written essay entrance examinations;
- (iv) performance in interviews and oral examinations and
- (v) performance on achievement and aptitude test.

As the first type of data is primarily the concern of the system of the secondary education, it need not be dealt with here. Attention may be directed toward the remaining four types, as they may be relevant for college and university admission in India. Subsequent paragraphs will consider each type of data separately.

Secondary School Performance

Assessment of secondary school performance has been generally carried out in two different ways in the countries reviewed. The first way consists of regular assessment of performance by teachers throughout the secondary school career, as in Japan and U.S.A., and the second way consists of a terminal examination, at least partly external, for a school leaving certificate as in U.K. In U.S.S.R., both types of assessment are employed and in all of the four countries, the assessment of secondary school performance is one of the important factors in deciding admission to institutions of higher

learning. To examine the effectiveness of secondary school assessment for predicting university performance, some research data published in U.K. and U.S.A. have been presented in the Table. The relationships between secondary school performance and performance at the end of the university degree course (generally four years later) have been reported as coefficients of correlation. The coefficients given in the Table may be interpreted as meaning that university performance is significantly related to secondary school results, but that predictions cannot be made with complete certainty. This interpretation is the same regardless of whether secondary school assessments were made by teachers' marks or terminal examinations. A number of factors may limit the effectiveness of secondary school assessments as a criterion for admission to colleges and universities. If secondary school assessment is made with the objective of evaluating potential university success, then it may be expected to be a more accurate admission criterion. However, generally secondary school assessments are carried out for purposes other than predicting university success. Also, if the subject-matter and methods of assessment in the secondary school are similar to those of the college and university, the secondary school assessment would be expected to be a better criterion than if subject-matter and methods of assessment are dissimilar. Other factors influencing the effectiveness of secondary school assessment as an admission criterion are the reliability of the secondary school assessments and the consistency of standards. Thus, if it is found that there is a high degree of agreement between examiners marking the same answer papers, and if standards remain relatively constant from year to year, between schools or secondary school examination boards, the secondary school assessment will be a more satisfactory criterion for admission purposes.

Essay Examinations

Supplementary data for admission purposes are collected using entrance examinations of the essay type in Japan, U.K. and U.S.S.R. Entrance examinations in Japan generally cover the following subjects: Japanese, mathematics, natural science, social studies and foreign languages. In U.K., entrance examinations are often conducted by individual colleges or departments, and generally

emphasize subject-matter considered to be basic to the proposed area of university study. Essay examinations are also conducted in the U.S.S.R., however students with exceptional performance in secondary school are exempted from taking them. The subject-matter of the examinations is generally determined by the courses for which they are held. Russian language and literature and one foreign language are commonly examined. For entrance to science, agricultural and engineering institutes, mathematics, physics and chemistry are also examined. (It may be noted parenthetically that while essay examinations are not widely employed in the U.S.A. for university admission purposes, they are used regularly in secondary schools, colleges and universities to evaluate the academic progress of students). In the absence of sufficient published data on the effectiveness of the essay type entrance examinations, conclusions regarding its adequacy should not be ventured. However, it is subject to certain limitations which may affect its value for admission purposes. Agreement and consistency between examiners with respect to standards of merit and distribution of marks, and the adequacy with which different topics belonging to the same subject are covered, may be noted. The cost of marking such examinations is also relevant in this context.

Interviews and Oral Examinations

Interviews, in which one applicant is met by one or more university staff members, are frequently employed in U.K. and also by private colleges and universities in U.S.A. Oral examinations, during which other applicants are also present, are widely used in U.S.S.R. In both interviews and oral examinations, spoken performance rather than written performance is evaluated. Some difficulties arise for this type of evaluation due to the absence of a permanent record of the actual performance. Maintaining the same standard over a series of applicants and objectivity of interviewer judgments are difficulties associated with assessment by interviews and oral examinations. This method may also be expensive in terms of interviewer time if the number of applicants is large. Lack of sufficient published data on the effectiveness of oral procedures for college and university entrance purposes does not permit their appraisal here.

Aptitude and Achievement Tests

Aptitude and achievement tests are widely employed in the U.S.A. for admission to private colleges and universities. Two hundred and forty colleges and universities are members of the College Entrance Examination Board, all of which require applicants for admission to take one or more of the College Board tests. Among these colleges and universities are California Institute of Technology, Columbia University, Cornell University, Harvard University, Massachusetts Institute of Technology, University of California, University of Chicago and Yale University. Historically, the methodology of aptitude and achievement tests can be traced to U.K., Germany and France. Sir Francis Galton in U.K. emphasized differences between individuals and developed statistical methods for their analysis. Methods for precise measurements of psychological characteristics were contributed by William Wundt in Germany. Alfred Binet in France developed standardised measures directed toward identifying children of differing levels of ability. Currently used aptitude and achievement tests have separate aims: aptitude tests seek to assess potential accomplishments or ability to learn, while achievement tests are designed to measure level of accomplishment or proficiency at present. Aptitude tests are used for admission purposes on the assumption that achievement tests alone may not be satisfactory predictors of university performance, partially because students' achievement will be affected by qualitative differences in schools attended, teachers and coverage of the subject-matter. Achievement tests are also employed to provide a comparison of the performance of applicants from different secondary schools.

As frequently employed in the U.S.A., aptitude and achievement tests are of the modern objective type, consisting of a large number of questions each of which is followed by several alternative answers. Candidates indicate their answers by choosing one of the alternatives for each question, and performance is indicated by a simple or weighted sum of the number of right answers. Characteristics of this type of aptitude and achievement tests include the possibility of a wide coverage of topics belonging to the same subject, and rapid and objective scoring. Considerably more time

and care are required to set the questions for this type of test, which in the case of large numbers of applicants may be offset by the saving in marking the answer papers. As these tests may consume less time than essay examinations, they permit collection of reliable and valid quantitative measures of aptitude and achievement in a number of different areas within a limited amount of time, which in turn may be appropriately used for placement as well as selection purposes. Some data illustrating the effectiveness of these types of tests are presented in the Table for U.K. and U.S.A. These data show significant correlations between aptitude test performance and later university results.

Concluding Remarks

The preceding paragraphs have been concerned with admission procedures considered singly. It may be interesting to know whether prediction of university success can be improved using a combination of two or more procedures. Illustrative data presented in the Table suggest that predictions can be improved if, in addition to secondary school assessments, aptitude test results are also used. If this combination is recommended for use in India, several questions might arise. One such question might concern the suitability of aptitude and achievement tests for use in India. Pilot studies on an experimental basis have been carried out to investigate this question in the Indian Statistical Institute. Where tests have been developed expressly for use in India, their correlations with academic performance are similar to those reported abroad. Another question might arise regarding the effectiveness of procedures currently used in India. To answer this question, data need to be collected regarding prediction of success in various courses by existing admission procedures, the reliability of predictions and their associated costs. These data might be collected for a random sample of colleges and universities in different geographical regions. It would also be possible to compare the effectiveness of procedures currently used in India with alternative procedures, such as those touched upon here, if both were tried out simultaneously on a pilot or experimental basis. This would call for a long term project to investigate the prediction of university performance in different subjects. Examples of long-term projects concerning predictors of

university success are being carried out in U.K. by the Nuffield Foundation and in U.S.A. by the National Merit Scholarship Corporation. A fixed admission procedure for use in India may not be desirable. On the contrary, the desirability of flexibility, in the interests of the universities and of potential students, and periodic evaluation of the effectiveness of the procedures being used, should be stressed. Further, the effective placement of students in priority fields should be considered. It has been reported that in U.S.S.R., enrolment in each subject is planned for the country as a whole and for separate institutions of higher learning. While this may not be recommended for colleges and universities in India, it is noteworthy that it implies flexibility of standards according to needs and that it emphasizes the role of higher education in national development. It is hoped that the points discussed here will assist in the formulation of principles and development of methods for placement and selection of students to be admitted to colleges and universities in India.

TABLE
*Correlations between Admission Data and Final University
Performance from Published Reports in U.K. and U.S.A.*

Admission Data.	Country.	Type of Assessment.		Correlation and significance.	Footnote.	Reference.
		Admission.	University.			
1	2	3	4	5	6	7
Secondary	U.K.	Scottish Senior Leaving Certificate.	Degree Class.	.14,.01	1	Could and M'Comisky (1958).
Secondary School Assessment.	U.K.	Scottish Senior Leaving Certificate.	Degree Class.	.34,.01	1	Dale (1952).
Secondary School Assessment.	U.S.A.	High School Record.	4 Year Cumulative Average.	.46,.01	2	French (1957).

1	2	3	4	5	6	7
Essay and Interview.	U.K.	Combined Essay and Interview Marks.	Degree Class.	.26, Not significant	3	Himmelweit and Summerfield (1951).
Aptitude Tests.	U.K.	Battery of Tests.	Degree Class.	.52,.01	3	Himmelweit and Summerfield (1951).
Aptitude Tests.	U.K.	Battery of Tests.	Total Final Marks.	.50,.01	3	Himmelweit and Summerfield (1951).
Aptitude Tests.	U.S.A.	Scholastic Aptitude Test (Verbal).	4 Year Cumulative Average.	.43,.01	2	French (1957).
Aptitude Tests.	U.S.A.	Scholastic Aptitude Test (Quantitative).	4 Year Cumulative Average.	.27,.01	2	French (1957).
Combined Criterion.	U.S.A.	High School Record and Scholastic Aptitude Test (Verbal).	4 year Cumulative Average.	.63,.01	3	French (1957).

1. Phi coefficient of correlation computed from published data.
2. Average product-moment correlation for 10 colleges and universities.
3. Multiple correlation.

Selected References

- Andersen, R. S. (1959). *Japan. Three Epochs of Modern Education*. Washington, U.S. Government Printing Office.
- Chauncey, H. and Frederiksen, N. (1951). The functions of measurement in educational placement. Chap. 4. in Lindquist, E. F. (Ed.) *Education Measurement*: Washington: American Council of Education.

- Cronbach, L. E. and Gleser, G. C. (1957). *Psychological tests personnel decisions*. Urbana: University of Illinois Press.
- Dale, R. R. (1952). The prognostic value of the University entrance examination, *British Journal of Educational Psychology* 22, 124-139.
- Dyer, H. S. and King, R. G. (1955). *College Board Scores*. Princeton, N. J.: College Entrance Examination Board.
- French, J. W. (1957). *Validation of the Scholastic Aptitude Test and new item types against four year academic criteria*. Princeton, N. J.: Educational Testing Service Research Bulletin RB-57-4.
- Gould, E. M. and M'Comisky, J. G. (1958). Attainment level on leaving certificate and academic performance at University. *British Journal of Educational Psychology* 28, 129-134.
- Himmelweit, H. T. and Summerfield, A. (1951) Student selection —an experimental investigation. II. *British Journal of Sociology*, II, 59-75.
- Korol, A. G. (1957). *Soviet education for science and technology*. Cambridge: Technology Press of Massachusetts Institute of Technology and New York: Jhon Wiley.
- Ministry of Education, Government of Japan (1959). *Education in Japan*. Japan: Government Printing Bureau.
- Thorndike, R. L. and Hagen, E. (1955). *Measurement and evaluation in Psychology and education*. New York: John Wiley.
- U.S. Department of Health, Education and Welfare (1957). *Education in the U.S.S.R.* Washington: U.S. Government Printing Office.
- Vernon, P. E. (1956). *The measurement of abilities*. 2nd Ed. London: University of London Press.

APPENDIX VII

The Tutorial System

Shri S. R. Dongerkery

The following conditions are necessary to make tutorials satisfactory :

(i) The topics for the tutorial essays should be carefully selected, after consultation among the members of the teaching staff of a Department. The topics should be selected so as to cover important branches of the subject.

(ii) In setting the tutorial essays the teacher should give reading assignments to the students sufficiently in advance to enable them to consult the books in the library and to write the essays at home.

(iii) The essays should be examined by the teacher who should correct and return them to the students before the topic is discussed in the class. The teacher should not give points for the essay, although he may discuss the topic in a general way before giving the assignments.

(iv) In the discussion the teacher should see that as many students as possible take part, either of their own accord or by way of answers to questions put by him. Except for explaining some difficult portion of the topic, or for correcting a wrong view expressed in the discussion, the teacher should allow the students to do most of the talking.

(v) A record should be kept of the essays written by each student, the attendance at discussion and the performance of the student both in writing the essay and in participating in the discussion.

(vi) Heads of departments should occasionally attend the tutorials of their junior colleagues, with a view to making suggestions, if any, for improving the conduct of the tutorials. Such suggestions should, however, be made outside the class, and never in the presence of the students. In order to avoid any possible misunderstanding on the part of the students about the object of such attendance, and also in order to enable junior colleagues to gain experience by attending the tutorials of their senior colleagues, the former should be asked to attend some of the tutorials of their senior colleagues.

(vii) Attendance by students at tutorials should be made compulsory. The University should consider the desirability of making the attendance at least at two-thirds of the tutorials in each subject compulsory before a student is sent up for the examination, or promoted from the junior to the senior B.A. class.

(viii) The number of students in a tutorial class should not ordinarily exceed 15.

(ix) The University should consider the possibility of giving credit, to the extent of 20 per cent of the marks, on the performance of students at the tutorials. This is, of course, a difficult problem, but some attempt must be made to solve it satisfactorily.

APPENDIX VIII

Grading Systems in Colleges and Universities of the United States and their practice in India.

Dr. Frank M. Fletcher and Prof. T. K. N. Menon

History and Introduction

All the early colleges and universities of the United States were established as private institutions, with complete independence and academic freedom. These institutions set the precedent for the establishment of public financed colleges and universities, which are now in the majority. The principles of independence and academic freedom were incorporated in the public financed colleges and universities, and these are highly prized and protected.

Along with the concept of academic freedom came the concept of freedom of the individual professor or instructor. This meant that, within the limits of good administration and institutional unity, the professor controlled the specific content of the courses taught and the determination of the proficiency of students. The professor, then, is the important and significant person. This does not mean that there is chaos in the curricula offered. On the contrary, professors spend much time in periodical meetings to discuss the various courses and the total curriculum, make modifications, and arrive at decisions by democratic means. In recent years various accrediting agencies and professional societies have played a significant role in providing guidance and helping to improve and maintain standards, but such actions have not, in any critical way, affected the academic freedom of the professor.

It should be emphasized that accrediting agencies only set minimum standards, and there is no pretence made that all colleges and

universities are of equal standard or that the graduates of all colleges and universities are equally good. Instead, each college and university must build its own reputation, and this largely results from the reputation attained by the graduates or alumni of the institution. There is large variation in the average calibre of students in various institutions. Hence grade averages are not equivalent from institution to institution. The honours graduate (one of the best) of one college may not be as good as a very poor graduate (low grade average) from some other college, in the extreme cases. This situation makes the faculty members in every college interested in maintaining as high a standard as is feasible in their own institution in order to turn out graduates who will improve the institutional reputation. Of course, the general goals of educational purpose varies. The whole development of the junior colleges is pertinent in this respect.

Organization of Courses

The academic year is universally about nine months. This period is divided in most colleges into two halves, called semesters. In some colleges there are three divisions, called quarters (the summer term may then be a fourth quarter). A course may last for only one term (quarter or semester) or it may continue over two or three terms. In either case, the term (quarter or semester) is a unit for grading purposes. That is, a final grade is awarded to each student for the course for that particular term. At the end of each term the student will receive a grade for each course taken during the term, and such grades are entered on the permanent cumulative record (transcript) of the student. These grades are averaged (as described later) and the student receives a grade average for that term as well as a cumulative grade average for all of the academic work completed in the college up to that time.

The grade averages (both for the term and cumulative) are checked each term and students not maintaining minimal averages will be dismissed. Dismissed students may apply for re-admission, and some, after careful screening, may be readmitted for further study on the condition that they maintain certain grade averages each term.

Grading Systems

Although there is some variation in grading systems by far the most common system is the following, sometimes with some slight variation:

<i>Grade</i>	<i>Meaning</i>	<i>Grade points</i>
A	Excellent	4
B	Good	3
C	Average	2
D	Poor	1
E	Fail	0

There is no satisfactory way of defining each grade in absolute terms — all grades after all are relatively based. The best students tend to be given an A grade while the poorest students tend to be failed with an E grade. Each professor develops his own concept of standards, and the distribution of grades over a period of time may vary considerably from professor to professor. Some are noted “tough” graders (giving few if any As and numerous Ds and Es). Others are easy graders (many As and few Es). Grading patterns tend to run quite consistently over a period of time. The overall average of grades given by a total faculty will remain very constant from year to year. A typical college student will receive from 12 and 18 grades per year given by 12 to 18 different professors. As a result the variation of grading among professors tends to be averaged out.

The grade points are assigned to grades in order to allow arithmetical averaging. Courses will vary from 2 (occasionally even 1) to 5 (sometimes more) hours credit. This depends on the number of hours or class periods per week. To compute an average for a term for a student, the grade points for each course are multiplied by the number of hours credit for the course, and the total for all the courses is divided by the number of hours giving the “grade point average” for the term. The same principle is applied in

computing the "cumulative grade point average" for the total academic work completed.

Each college or university has certain specific regulations regarding the grade point average that must be maintained by a student to stay in the college or university, and a minimal average that must be attained over the four or five year period to graduate and be awarded a degree.

Basis for Grades

There are no specified regulations regarding just how a professor arrives at a grade for a student. This is the prerogative of the professor. Except for high level and advanced courses there is, however, a rather common pattern. Examinations usually count the most, but the number of tests or examinations will vary considerably depending on the professor and the nature of the subject being taught. Almost universally a final examination is given at the end of the term and counts for one fourth or a bit more for the final grade. The last week of any term is given over to examinations, and is called the final examinations week. Tests or examinations given during the term may vary from one at the middle (called a mid-term exam) to one every two weeks, one every week, and even now and then a short quiz each day. In some courses the final grades may be based almost entirely on the examination results, but most frequently other grades or factors are also taken into account. Such factors include class discussion, term papers or projects laboratory work, special short term papers or projects, and a variety of other things. In general, although there is a good deal of uniformity, there is also a great deal of variation.

Validity of Grades and Problems

Considerable attention has been given to the validity, reliability, and variation of grades. It is well recognized as based on much evidence that the validity (and even reliability) of grades by no means approached the desired level. Of course, it must be recognized that due to lack of adequate criteria it is not possible to

accurately determine the validity of grades. If it were possible to apply more stringent control and training on professors, the grades could be improved, but the "sacred" academic freedom makes this most difficult. At the same time, much is being done to aid professors in developing better examinations, etc. which does not help. In spite of the difficulties and shortcomings, the total system works surprisingly well-in fact, very well. The value of freedom in instruction and the advantages to society of variation in instruction are so important that it is certain they will not be disturbed to enable a slight increase in the validity and reliability of grades. After all, grades are definitely secondary in good education Society profits, not from the grade a student receives, but from what a student gains to enable him to contribute to society at a later date.

There has always been, and always will be pressure put on professors regarding grades. This matter has been well solved in the United States by the fact that professors stand together on this point. If one professor is attacked everybody else goes to his aid. In unity there is strength. As a result the public has well learnt that it does little good to try and apply pressure. No doubt, many still try it, but get nowhere.

Grade-Credit System in Indian Universities

The grade-credit system has not been attempted by Indian Universities. An exception to this is the M. S. University of Baroda which has been pursuing this practice in a few of its Faculties. The best example of the working of the grade-credit system with adaptations to suit Indian conditions is to be found in the Faculty of Home Science of the University. The Faculty follows the system to assess and evaluate the work of its students in theoretical as well as practical work.

Credit

All the courses given in the Faculty are assigned a specific number of credits. One credit in a course means two lectures each of 45 minutes' duration or a laboratory practical of 90 minutes duration

per week in the subject throughout the term. It also includes 180 minutes of home and library work.

Grade

A student's sessional work is assessed on the basis of grades. Every letter-grade carries both qualitative and quantitative values as shown below :

<i>Grade</i>	<i>Qualitative Values</i>	<i>Quantitative Values</i>
A	Excellent (First Class)	3 points
B	Good (Second Class)	2 points
C	Average (Third Class)	1 point
F	Poor (Failure)	0 point

The Procedures for Grading

Every teacher teaching a course maintains a Class Record Card for the subject in respect of each student taking the course. The student's performance during a term in the subject in areas of periodical tests, term papers, reports, classroom work and special projects are graded on the above mentioned four point scale and the grades earned by the student in each area of the sessional work is entered in her Record Card. All these grades are averaged and count as 50 per cent of the total grade for the term. At the end of each term a terminal examination is conducted, which covers the whole course for the term. Grade for this examination and the over-all grade for the sessional work during the term are added and the average of the two is calculated for awarding the final grade of the term in the course.

In determining a grade, the following points are considered:

1. The student's achievement;
2. The progress she shows over her previous work;
3. The quality of her practical work;

4. The skills, management ability and judgement displayed by her;
5. The extent to which she has made use of the opportunities given to her to improve her grade; and
6. Her performance in relation to the total performance of the group taking the course.

Assigning of Over-all Grade

The over-all grade assigned to a student at the end of a year is based on the average of total points calculated on the basis of (i) the number of courses taken by her, (ii) the number of credit-points carried by each course taken by her and (iii) the quality points of the grade earned by her in the course during the year. The student's over-all grade is the average of values of these three factors. The following example of the achievement of a student will illustrate what is described above:

<i>Course</i>	<i>Credits</i>	<i>Grade earned by the student</i>	<i>Quality points of the Grade</i>	<i>Total Points</i>
No. 1	6	B	2	$6 \times 2 = 12$
No. 2	2	A	3	$2 \times 3 = 6$
No. 3	4	A	3	$4 \times 3 = 12$
No. 4	6	A	3	$6 \times 3 = 18$
No. 5	6	C	1	$6 \times 1 = 6$
No. 6	4	B	2	$4 \times 2 = 8$

Total points = 62

$$\text{Average} = \frac{\text{Total Points}}{\text{Total Credits}} = \frac{62}{28} = 2.2$$

* Over-all Grade = B

Standard for Promotion

A student, for being promoted, should pass in more than half the subjects taken by her during the year and maintain an over-all C grade.

The grade credit system has worked quite well. It has made the student work diligently and regularly. Opportunities are provided to her to improve her grade. And the emphasis is not on failing a weak student but on helping her to come up to the level of attainment she is normally capable of. The satisfactory working of the grade-credit system in the Faculty has not only improved its testing and evaluation practices, but it has also resulted in the improvement of its syllabus followed for the study of different subjects.

APPENDIX IX

Grading System in Japan

Prof. Samuel Mathai

The work to be done by a student is measured by “credits”. A typical illustration of the Credit system may be given from the requirements of one University (Waseda).

One unit of credit is defined as the credit given to a lecture course or seminar of one hour per week for a term of 15 weeks. For laboratory work, one unit of credit is given for 3 hours of laboratory work per week for a term of 15 weeks. For exercises in mathematics and foreign languages, one unit of credit is given for 2 hours of exercise per week for a term of 15 weeks.

For a bachelor’s degree, a student must have taken 132 units of credit in accordance with the School or departmental requirements.

(a) Of these 132 units, 36 units (4 units for each course) must be taken in subjects designated as those of general education. Three courses (12 units) must be chosen from each of the three categories of subjects offered by the School in which the student is enrolled. These categories are Humanities, Social Sciences, and Natural Sciences.

(b) 12 Units of credit must be taken from among courses in two foreign languages during the first two years at the University.

(c) 4 units must be in physical education, 2 for theory and 2 for practice.

(d) Credits in specialized subjects must be taken as follows :

Schools.	Departments.	Specialized subjects		Total units of credit.
		Compulsory Units.	Elective Units.	
1	2	3	4	5
Political Science and Economics.	Pol. Science	48	40	88
	Economics.	"	"	"
	Journalism.	64	28	92
	Local Govt.	44	44	88
" (Evening Division).		40	40	80
Law.		52	36	88
" (Evening Division).		52	28	80
Literature.		48	32	80
" (Evening Division).		40	40	80
Education.		52	32	84
Commerce.			76	76
„ (Evening Division)			"	"
Science & Engineering.				84
	Mechanical Engg.	62	22	
	Electrical Engg.	72	12	
	Mining Engg.	51	33	
	Architecture.	60	24	
	Applied Chemistry.	78	6	
	Metallurgical Engineering.	66	18	
	Electrical Communication.	50	34	
	Industrial Management.	43	41	
	Civil Engg.	58	26	

1	2	3	4	5
Science and Engineering (Evening Division).	Applied Physics.	72	12	
	Mathematics.	68	16	84
	Mechanical Engineering.			
	Electrical Engineering.	54	30	
	Architecture.	62	22	
	Civil Engineering.	52	32	

While the unit of credit is more or less the same in all the Universities, the number of units required varies slightly from University to University. This system of credits replaces the examination system for all practical purposes. Universities do have entrance examinations and other tests, but it is the credits that determine a student's academic standing and his right to proceed with his education.

APPENDIX X

*The Combination of Internal Assessments With Examination Results**

Dr. H. J. Taylor

It is generally agreed that in evaluating a student's performance his record in school or college should carry some weight. The evaluation should not depend entirely on the result of a single examination. This note is therefore concerned with the problem of combining internal assessments made in the school or college with the results of final examinations.

An examination mark is presumably a measure of some characteristic of the candidate, which we may call X. This X is usually a complex and undefined mixture of many things—knowledge, memory, intelligence, reasoning power, power of expression, handwriting, etc. The fact that successive examinations of the same candidates often show high correlations indicates that X, in spite of its vagueness, may remain fairly constant in examinations of the same type.

An internal assessment, however, depends on cumulative scores over a considerable period of time. It is based on tests which may differ radically from those which enter into a conventional examination. It is thus by no means obvious that the internal assessment and the examination are measuring the same thing. The examination measures X, but the internal assessment may measure some quite different characteristic Y.

*Reprinted from the *Journal of Education & Psychology*, M. S. University, Baroda, Vol. XIX No. 4, January 1962.

If X and Y are the same, the problem of combining the measures involves mathematical concepts only. If X and Y are different, the problem is no longer wholly mathematical. It is conceivable, for instance, that X is highly correlated with memory, whereas Y depends on diligence. The extent to which a good memory should be allowed to outweigh a lack of diligence is a matter for personal judgment. Until that judgment is made statistical methods do not apply.

In the present discussion we shall assume that X and Y are identical, so that the internal assessment scores and the examination marks may be regarded as different measures of the same thing.

For convenience, let us denote the internal assessment by I.A., and the examination by E . For a given set of candidates, I.A. and E are each presented in the form of a distribution of marks. The two sets of marks may have the same maximum, but they will usually show a difference of scale. This implies that the two sets will usually have different means and different standard deviations.

Before combining marks they should be scaled to the same mean and standard deviation. If this is not done large errors will be introduced, and indeed the failure to use scaling techniques is one of the major weaknesses of current examination methods. The necessity of scaling, however, arises in all examination procedures, and is not peculiar to the problem considered in this note. We shall therefore assume, in what follows, that the marks of I.A. and E . have already been correctly scaled to the same mean and standard deviation.

Let x denote the mark obtained by a candidate in E . This mark is an estimate of the character X which the examination is designed to measure. It is important to know how accurate this estimate is; in other words, what is the precision of the mark x ? If the candidate were examined a hundred times he would not get a hundred identical marks, the marks would be scattered more or less widely round a mean value. If they were closely grouped we would conclude that the examination measures X rather precisely, but if they were widely scattered we would conclude that the measurement

is rough and inaccurate. The spread of the marks would be a measure of the precision of the examining technique.

The spread is suitably defined by the 'probable error', or p.e. This is a range about the mean so chosen that half the marks lie within it and half outside; we denote it by $\pm p$. The mark must now be represented not simply by x , but by $x \pm p$, where p denotes the imprecision attaching to the assigned mark x .

It is not practicable to determine p by conducting a large number of similar examinations, but there are other ways of estimating it. In what follows it will be shown that a knowledge of p is required before the results of I.A. and E. can be combined with confidence.

The two measures given by E. and I. A. are denoted respectively by $x \pm p$ and $y \pm q$
or by $x \pm ax$ and $y \pm by$
where a and b are the relative p.e.'s, namely the ratios of p and q to x and y respectively.

The combined mark is given by $M = x + y$

By the usual theory, the p.e. of M is given by

$$m = \sqrt{a^2x^2 + b^2y^2}$$

The relative p.e. is $m/M = c$, which is given by

$$c = \sqrt{a^2x^2 + b^2y^2} / (x + y)$$

$$a = \sqrt{1 + b^2y^2/a^2x^2} / (1 + y/x)$$

Writing $b/a = r$, and $y/x = s$

$$c/a = \sqrt{1 + r^2s^2} / (1 + s) \quad \dots (1)$$

This equation gives c/a , which is the uncertainty of the result M in terms of the uncertainty of the examination mark x . Since c/a may be greater than unity, it is possible for the combined mark M to be more uncertain than x alone. We do not necessarily improve the evaluation, therefore, by adding I.A. to E.

We shall assume for simplicity that E. is more accurate than I.A. or that a is less than b . The following two cases are of interest:

Problem 1: For a given r , what value of s makes $c/a = 1$?

By putting $c/a = 1$ in equ. (1) we find

$$s = 2/(r^2 - 1) \quad \dots(2)$$

Table I gives corresponding values of r and s according to this equation.

Table I

$r =$	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.8	3.2	3.6	4.0
$s =$	4.55	2.08	1.28	0.89	0.67	0.52	0.42	0.29	0.22	0.17	0.13

Since it was assumed that the marks of E. and I.A. were originally given in the same scale, the quantity s is the relative weight assigned to y when the marks are combined. (Thus, if $s=0.5$, the y marks are divided by 2 before being added to the x marks). If, for any given value of r , y is given a weight s greater than that shown in Table I, the combined measurement M will be more uncertain than x alone.

Problem 2: For a given r , what value of s gives a minimum value of c/a ?

This problem is solved by differentiating equ. (1) with respect to s , and equating to zero. The result is

$$s = 1/r^2 \quad \dots(3)$$

This relation is exhibited in Table II:

Table II

$r =$	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.8	3.2	3.6	4.0
$s =$	0.69	0.51	0.39	0.31	0.25	0.21	0.17	0.13	0.10	0.08	0.06

For these values of s , c/a is a minimum, and the combined result M has the maximum accuracy. Since the values in Table II are all smaller than the corresponding values in Table I, the combined result M is in this case always more precise than x alone.

Discussion. The accuracy of the combined result M is seen to depend on the relative weight s which is given to the measurement y , as compared with x . The proper value of s to be used is a function of r , the relative accuracy of the two sets of marks. We thus reach the following conclusion :

Before we can combine I.A. marks with E. marks, it is necessary to know the relative accuracy of the two sets of marks.

In current practice no attempt is made to determine r , and the value of s is laid down by rule of thumb. A regulation is made, for example, that 25 per cent weight shall be given to I.A., and 75 per cent to E. But this procedure is full of pitfalls. The relative weights are assigned, presumably, on the basis of a general feeling that I.A. is less reliable than E. If it is felt that E. is three times as accurate as I.A., the E. marks should be given three times the weight. But Table 1 shows that this is fallacious. If the relative accuracy does in fact correspond to $r = 3$, the weighting factor s should not be $1/3$, but $1/4$, if the final result is not to be worse than the examination mark taken alone. But putting $s = 1/3$, as the naive reasoning suggests, the final result becomes worse, in point of accuracy, by some 6 per cent. Table II shows that for the best results s should be much smaller; for $r = 3$ the best value is $s = 1/9$.

The conclusions to which the analysis leads may be summarised as follows :

(1) Unless the relative accuracy of the two measurements can be estimated they should not be combined at all. An *ad hoc* procedure is liable to make the reliability of the result worse instead of better. Whichever of the two appears, on general grounds, the more dependable, should be used alone and the other ignored.

(2) To make the combination of I.A. and E. acceptable in practice, attention should be concentrated on methods of estimating the accuracy of examination marks. A possible method is that of the divided examination (see appendix).

(3) Whenever the relative accuracy can be estimated, weightage

must be given in accordance with Table II. It is fallacious to assign relative weights in advance by some arbitrary rule.

SUMMARY

The weighted mean of a final examination mark and an internal assessment is not necessarily more accurate than the examination mark alone, and may often be less accurate. The paper shows that the optimum weighting factor is a function of the relative accuracy, r , of the two sets of marks, and it is therefore important to estimate r before the marks are combined. The attached paper discusses the divided-examination method of estimating the accuracy of marks.

The Estimation of Accuracy by the Divided-Examination Method.

Suppose a group of candidates to be given two examinations which are equivalent, in the sense that they test the same thing in the same way. The questions would be equally numerous, be of the same standard, offer the same degree of choice, and cover the same field. If examinations were ideal tools of intellectual measurement the two results should be identical. We get some approach to this state of affairs in intelligence testing, where the measured I.Q. is much the same whatever battery of tests is used.

In practice the two results would differ, partly because the character X which the examination measures is not well defined and partly because the conventional examination is not a precise tool. The differences in the two sets of marks provide a measure of the uncertainty or imprecision of the marking.

The best way of realising two equivalent examinations is to divide a single examination into two parts. Suppose we have a large number of questions, covering the whole field, which are presented to the candidates in random order. After marking, the odd-numbered questions are taken as one group, and the even-numbered questions as another. These two groups may then be considered as equivalent examinations which may be compared.

In the conventional examination it is common to have, say, four papers in a given subject, but each only covers a limited part of the field. Thus in Physics the papers may be: I, Mechanics; II, Heat; III, Optics; IV, Electricity. This examination cannot be divided into two equivalent parts. To make the division possible the questions should be mixed, so that a single paper may contain questions from all the branches. The actual questions could be the same as before, the important thing is the random distribution. We could then regard Papers I and IV as an examination group which would be closely equivalent to Papers II and III considered as another group.

To design a large public examination in this way would doubtless demand considerable thought. In particular the effect of offering a choice of questions would need to be investigated. But it is certainly possible to overcome the difficulties, without impairing the validity of the total examination. The division of an internal assessment should present fewer difficulties, since such assessments are usually derived from a large number of separate tests.

Let E_1 and E_2 denote two equivalent groups obtained from an examination E . The group E_1 provides a set of marks x_1, x_2, x_3, \dots , and E_2 provides a set x'_1, x'_2, x'_3, \dots . Subtraction gives a set of differences

$$(x_1 - x'_1), (x_2 - x'_2), (x_3 - x'_3), \dots$$

If the marking were completely accurate all these differences would be zero, and their actual magnitude provides a measure of the inaccuracy. One may take either the mean absolute difference (without regard to sign), or the root-mean-square difference. Assuming that the marking errors are distributed normally, either of these is proportional to the p.e. of the marking. The root-mean-square is slightly more reliable than the mean, but probably not sufficiently so to justify the extra labour involved in calculating it. The mean absolute difference has the advantage of being easily determined.

This procedure involves the assumption that the error in marking does not depend on the magnitude of the mark. This assumption

is questionable. It is perhaps more likely that the error is proportional to the mark, i.e. if a mark of 50 is liable to an error of ± 10 , then a mark of 25 would be liable to an error of ± 5 . On this basis it would be preferable to use not the actual differences, but the proportionate differences, obtained by dividing the difference by the mean mark. If these quantities are denoted by e , we shall have

$$e_1 = 2(x_1 - x_1') / (x_1 + x_1')$$

The absolute mean value of e will then provide the required measure of the inaccuracy of the marking.

The application of these ideas to conventional examinations, both theoretically and experimentally, offers a fruitful field of research. This research should precede the large-scale introduction of internal assessments as supplements to conventional examinations.

APPENDIX XI

The Marks of Examiners

Dr. H. J. Taylor

In the college examinations for which I am responsible it has been the practice for examiners to submit their original mark-sheets without any departmental adjustments. I have used these sheets to study the way in which examiners actually distribute their marks. The study has ranged over the four College years, a wide range of subjects, fifty or more examiners, and a dozen successive examinations. The results have an obvious bearing on the question of the reliability and validity of marking, and on the procedures for scaling and combining marks. The present paper outlines some of the main points which have emerged from these studies.

From each mark-sheet the distribution is first plotted on squared paper. The y-axis is graduated from 0 to 100, and each mark from the sheet is recorded by placing a dot against the appropriate graduation. Several marks of the same value thus appear as a horizontal row of dots. In this way a histogram of the marks is built up, which gives a clear picture of the distribution. This can be done quickly, it takes only a few minutes to plot a mark-sheet with 100 entries. The points corresponding to the median, and to the upper and lower quartiles, are marked on the graph. The distribution is then described by the notation 'M (p, q)' where M is the median, p the range from M to the upper quartile, and q the range from M to the lower quartile.

It would of course be possible, by purely numerical methods, to calculate the mean and standard deviation (and other statistics) of the distribution, but the graphical method offers certain advantages for data of this kind. The accumulation of marks at particular

scores, for instance, is seen at a glance in the histogram, but might not be evident otherwise. Moreover, without modern computing facilities, the time taken by numerical calculations would greatly reduce the amount of material one could deal with. The simple description 'M (p, q)' is not only the easiest to obtain, but is probably the most useful for the purpose in hand.

About half the distributions show significant departures from symmetry. Positive and negative skewness both occur, and examiners are not always consistent in the type of curve they produce. In extreme cases p may be twice as large as q, or *vice-versa*. In marking equivalent sets of papers (*i.e.*, the answers of two random groups of students to the same question paper) one examiner may produce a positive and another a negative skew curve. Even if one assumes that the marking is reliable (in the sense that the examiner, re-marking the same papers, would produce the same curve) these departures from normality raise serious problems when the marks have to be scaled to a common standard.

Some examiners are apt to produce a J-shaped curve. Usually, but not always, the mode occurs at the pass mark. For example, in a Bengali vernacular paper the examiner produced the distribution '39 (4, 3)' but the mode occurred at 36. This mark was given to 19 scripts out of 96, a number much too large to be attributed to chance. An examiner in Chemistry produced the distribution '39 (11, 9)' with 16 papers out of 126 receiving the pass mark 30. This again can scarcely be a chance fluctuation, as calculation shows. The effect seems to indicate an unwillingness or inability to assess a paper accurately at or near the pass level. The student is given the benefit of the doubt, with a resulting accumulation at the pass mark.

The interquartile range $p+q$ may be taken as a measure of the standard deviation. If the curve were truly normal, this range would indeed be $4/3$ of the standard deviation. The standard deviations of the distributions, estimated in this way, vary very widely, not only between different examiners on the same paper, but between the same examiner's curves on different occasions, and, most noticeably of all, between different subjects. Mathematics shows the largest standard deviation, followed by Logic, Chemistry,

and Physics, typical values being 19 for Mathematics and 12 for Physics. Most Arts subjects show much lower standard deviations, typical values being Civics 8, Bengali 7, English 6. When one considers the fallacy involved in adding marks of different standard deviations, it is disquieting to note how large the differences actually are, and to remember that University tabulators take no notice whatever of standard deviations.

Large groups of answer scripts are frequently divided into sub-groups and sent to different examiners. By allotting roll numbers at random, or otherwise mixing the candidates, one can ensure that the sub-groups are statistically equivalent. Care has been taken to do this in all the present studies, so that any differences in the mark-distributions can be attributed to the examiners and not to the candidates. Such differences are often surprisingly large. On an English paper examiner A produced the distribution '51 (5, 3)' while examiner B on the same paper, marking another sub-group, produced the distribution '31 (3, 4)'. With one exception all the marks given by B were lower than any mark given by A. This means, effectively, that the score of any candidate was determined almost entirely by the accident of his roll number, and not by his performance.

The above is admittedly an extreme case, but it is common to find quite large differences, even in the science subjects. A first-year paper in Biology, involving 210 candidates was marked by two pairs of examiners, A and B taking 110 scripts, C and D the remaining 100. The median mark was $66\frac{1}{2}$ for A and B, and 42 for C and D. The true means were also calculated, the values being respectively 64.9 and 40.4. The difference in the average mark (whether one uses the median or the mean) is thus $24\frac{1}{2}$. With A and B, 88 per cent of the candidates got 55 marks or more; with C and D on the other hand 95 per cent of the candidates got less than 55 marks. The difference is even more remarkable in that each total mark is found by combining the separate marks of two examiners, a procedure which would tend in general to smooth out differences in the standards of marking. The examiners were very unwilling to admit even the possibility of such large differences until the analysis was put before them.

With some examiners the evaluation of a paper is affected by the impression they have received from the immediately preceding papers. This has been studied in detail in only one case, where the examiner produced a J-curve in which the pass mark 30 was given to an unduly large number of candidates. Many of these 30's occurred in successive pairs on the mark sheets, and calculation showed that the frequency of such pairs was much greater than could be attributed to chance. This indicates a tendency for the estimate formed on one paper to be carried over to the succeeding paper. This effect may well be fairly common, and it is one which merits closer study.

A study of correlations between sets of marks throws light on the validity of the marking. The following case is of particular interest. A group of some 120 third-year students was examined in theoretical and practical Physics and Chemistry. The following correlation coefficients (r) were calculated, together with their standard deviations (s):

Physics theory vs. Physics practical	$r=0.14$	$s=0.09$
Chem. practical vs. Physics practical	$r=0.14$	$s=0.10$
Physics theory vs. Chemistry theory	$r=0.30$	$s=0.09$

The last correlation was affected by the presence of two very poor candidates who scored zero in both examinations, and who should probably not have been in the class at all. If these are omitted the value of r falls to 0.20, which is barely significant.

Although these correlation coefficients are slightly larger than their standard deviations, they do not differ significantly from zero. These examinations were conducted with every care by the usual methods and by experienced examiners. If the marks are valid measures of ability one must conclude that there is no appreciable connection between practical and theoretical ability in Physics, none between the two practical abilities, and none between the two theoretical abilities. Such a conclusion is surely unacceptable, and the only alternative is that the marks are not valid measures at

all. They can only be regarded, in fact, as sets of random numbers.

These results clearly need to be extended by further studies, but they are already sufficient to throw grave doubts on the validity of our present examinations.

The present methods of combining and tabulating marks in University examinations, without reference to the distribution curves, are highly unsatisfactory. It is true that moderators are appointed to adjust the differences between the standards of the various examiners, but the methods they use do not have the effect of bringing the marks to a common scale. These methods are not only inadequate, but sometimes wholly fallacious, and may often increase the discrepancies they are intended to remove. Marks have an inherent uncertainty, which can only be reduced by radical changes in the whole technique of examining. But even with the present techniques there is surely no excuse for adding new uncertainties, after the examination is over, by faulty numerical methods.

APPENDIX XII

Classifying Students by Examination

Dr. H. J. Taylor

In a former note ('The Marks of Examiners') an account was given of the way in which examiners' marks are actually distributed. Some implications were pointed out, in particular the urgency of adopting simple scaling techniques before combining the marks of different examiners. The present note is concerned with the following question: To what extent can we rely on the classification of students in four groups (I, II, Pass and Fail) on the results of an examination?

There is a widespread feeling that these classifications are extremely unreliable. Any teacher can cite numerous instances in which a student is put into a different group from that which previous experience would suggest as appropriate. Very good students often get poor results, and conversely poor students get good results. But general impressions are vague, and may be disputed. What is needed is some more definite estimate of the extent to which misclassification can actually occur. An illustrative example has therefore been worked out in detail.

We assume first that there is a 'True Mark', which is the mark a paper would receive from an examiner whose marking is free from error. We then assume that the actual marks awarded are subject to error, so that the examiners' marks may be a little more or less (in extreme cases considerably more or less) than the true mark.

(Whether we can indeed speak of a 'true mark' at all is admittedly a difficult problem, particularly with essay type question papers

But it is generally assumed that there is, for each paper, some mark which truly represents the candidate's performance. We accept this assumption for the purpose of the present discussion).

Let us consider a group of 1000 students, whose true marks are distributed in such a way that the average mark is 45, and the standard deviation is 15. Broadly speaking, this means that half the students will have marks between 35 and 55, the others will fall above and below these limits, thinning out rapidly as one approaches extremely high or extremely low marks. We shall further assume that *Pass* is secured by a mark of 30, *II class* by a mark of 48, and *I class* by a mark of 60. This agrees with customary practice. Calculation then shows that on the basis of the true marks the students will be grouped as follows:

I Class	167 Students
II Class	267 Students
Pass Class	416 Students
Fail	150 Students

This distribution represents reasonably well what actually occurs in many typical examinations.

We next assume that the uncertainty of the examiner's estimate has a standard deviation of 5 marks. This means that his estimate is just as likely to be too high as too low, but that on about half the papers he will not make an error greater than 3 marks either way. Some errors will of course be larger, but the larger the error the less often it occurs. Only in rare and exceptional cases would the error exceed, say, 12 marks either way. An error of as much as 15 marks would hardly ever occur.

It must be emphasised that this represents extremely good marking. Of very few examiners could it be said that their estimates, on the average, are correct to '3'. We are not considering in this example the effect of wildly inaccurate marking, but of very reliable marking, marking which is probably much better than is usual in University examinations.

Consider now the effect of the examiner's errors. Students whose true mark is, say, 50, are liable to get marks above or below that value. Some may thus fall below 48 and lose their II class. Very occasionally, if the examiners' error were exceptionally large, such a student might be raised to the I class. It is possible to work out numerically, on the given assumptions, just how many of the students would be moved up and down in this way, and to what extent. The results of the calculation are as follows:

I Class students placed in the II class	26
II Class students placed in the I class	39
II Class students placed in the Pass Class	49
Pass Class students placed in the II Class.	52
Pass Class Students who fail	37
Failures who are placed in the Pass class	25

The number of students in the four groups are as follows :

	<i>True marking</i>	<i>Examiners' marking</i>
I Class	167	180
II Class	267	257
Pass Class	416	401
Fail	150	162

The general effect of the errors is to spread the marks more widely, so that we get both more I classes and more failures. Of 1000 students, 112 are put down to the next lower group, and 116 raised to the next higher. Altogether 228, or almost one quarter of the students, are wrongly placed.

These figures are very striking, all the more so when we reflect that the situation in actual examinations must be worse than this. We have assumed very good marking, with marks normally distributed. Actual examinations have many uncertainties over and above the error of estimation, and the previous note (The Marks of Examiners) showed that the marking of examiner is often not even approximately 'normal'. Every additional source of uncertainty must reduce the correlation between the true classification and that

provided by the examination. The example shows that even in a very good examination a quarter of the students may be wrongly classified and the fraction may well rise to one half for examinations as conducted in practice.

In the above example only the Pass students run the risk of failing as a result of the examiner's errors. We may call this 'failing by accident' since the student really deserves to pass. In this example 9 per cent of the Pass students fail by accident. In real examinations the percentage may well be greater, but could scarcely be less. In most examinations students must pass independently in a number of papers. What then is the chance of a pass student failing by accident on the whole examination?

To simplify the problem, suppose the students to have much the same ability in all subjects. On their true marks they would secure a Pass in every subject, and would therefore pass the whole examination. Since 9 per cent, however, fail by accident on each paper, the chance of passing in one paper is 91 per cent. From this we find the chance of passing simultaneously in several papers, as follows:

Number of Papers	2	3	4	5	6	7	8	9	10
Proportion of students who pass completely (in percentage)	83	75	69	62	57	52	47	43	39

It follows that although the candidates are by hypothesis above the pass level in all subjects, large numbers will fail 'by accident' in any examination where independent passing in several papers is required. With ten papers 61 per cent of the students will fail.

These considerations may provide some justification for the low pass mark in Indian University examinations. One might argue as follows: The true pass mark should be at least 40, but most of those who fail by accident at the 40 level will still get marks above 30. If therefore we make the pass mark 30, we ensure that almost everybody passes who really deserves to do so, at the cost of letting

through a large number of others. In other words, very few fail who ought to pass, but many pass who ought to fail. This does to some extent reflect the present position, and the argument is not without force. But we need a better method, it is not good enough to sweep up the wheat and chaff together as the only way of collecting the wheat. We must learn how to sift one from the other.

What conclusions can be drawn from these considerations? One may suggest at least the following three:

- (1) Marking must become more reliable. This means a larger element of objective testing in examinations, and reducing by this and other means the vagaries of the individual examiner's judgment.
- (2) The excessive deference paid to I class results is not justified. The published I class list contains many II class students, and possibly a few of even lower rank. Conversely many I class students will be found in the published II class list. It follows that in the selection of students for appointments, for higher studies or for scholarships etc., where some minimum standard is set (say I or II class) consideration should also be given to those who are in the next lower class. Otherwise some of the best people will certainly be missed.
- (3) The requirement that students must pass separately in a large number of papers should be abandoned. The average performance over a number of papers is the best single criterion of a student's performance. By averaging a number of results the effect of examiners' errors is appreciably reduced. The average mark on 9 papers (provided the marks have been properly scaled) is a more accurate estimate of a student's performance than any single mark — to be precise it is three times as accurate. The criterion for passing an examination should be the attainment of some stated average mark over a fairly large number of papers, the papers being marked independently by different examiners and these marks reduced to a common scale before the average is taken.

APPENDIX XIII

Choice and Chance in Conventional Examinations

Dr. H. J. Taylor

Introduction

In a conventional examination the candidate is offered only a small number of questions. These may be regarded as a sample drawn from a much larger number of possible questions. Since the student's knowledge is not perfect, an element of chance is involved in presenting the questions. If he is acquainted with one half of the field of knowledge covered by the examination, an undue proportion of questions may by chance relate to the other half, so that the candidate finds the examination too difficult. Conversely he may find it too easy. This element of chance in the random presentation of questions is well known and generally recognised. It decreases the reliability of the examination, regarded as a measure of candidate's knowledge.

It is usually assumed that the uncertainties introduced by chance are compensated, partly at least, by providing a choice of questions. Thus a candidate may be offered nine questions of which he must choose five. The assumption that the reliability of an examination is necessarily improved by this device is fallacious. Not only is the reliability not increased, but the provision of choice has the effect of biasing all the scores. The purpose of this note is to point this out by working out a particular example in detail.

Assumptions

To make the example definite and amenable to calculation the following assumptions are made. While this example represents

a simplified model of the real examination, it corresponds broadly to the situation as we find it in conventional examinations.

1. We assume that the examination is intended to test a student's knowledge over a wide field.
2. We assume that the field can be divided into a large number of items, more or less equivalent, and that each question tests the student's acquaintance with a single item.
3. We assume that one third of the whole field is well known to the student, one third only sketchily known, and one third not known at all.
4. We assume that the questions set are a fair sample from the field, so that for a question taken at random there is a probability of $1/3$ that the student can answer it well, a probability of $1/3$ that he can only give a poor answer, and a probability of $1/3$ that he cannot answer it at all.
5. We assume that the examiner sets nine questions, and that the candidate is only required to answer five.
6. We assume that the student always chooses a question on a well-known item in preference to one on a badly known item. (These questions may be named Easy and Hard respectively). He does not attempt those which are outside his knowledge.
7. We assume that the maximum mark for each question is 3, that a good answer on the average gets 2, and that a poor answer gets 1.

Analysis of the problem

On the above assumptions, it is clear that on the average, out of every three questions, the student should get 2 marks on one, 1 mark on another, and 0 on the third. Out of a very large number of questions presented without any choice he would thus average one mark per question. On the questions therefore, if he is not able to choose, his total mark, on the average, should be 5. We may regard this mark as the correct measure, under the given conditions, of his knowledge.

The question now arises, how will his total mark be affected by the circumstance that he is presented with a sample of nine questions, and is free to choose five of them in the most favourable manner?

There are three types of question of equal likelihood, Easy (E), Hard (H), and Impossible (I). A sample of 9 questions can be made up of E, H, and I in 55 different ways, and for each way we may calculate the probability of its occurrence. By assumption 6 each of these 55 groups of nine implies a definite choice of five and for this choice we can write down the total mark by assumption 7.

If we have a large number of candidates whose knowledge is distributed over the course in all possible ways (in conformity with assumption 3), then the number of candidates getting any particular mark will be proportional to the sum of the probabilities for that mark. Thus we may find what proportion of candidates get 1 mark, 2 marks, 3 marks ... etc.

The Effect of Choice

We may make an entirely analogous calculation in which the number of questions on the paper is only five, so that the candidates get no choice. Again we may find the proportion of candidates who secure each particular mark. The results of both calculations are placed side by side for comparison in the table below :

Table II

Total mark	Proportion of candidates getting the mark	
	(a) 9-question paper	(b) 5-question paper
1	2	3
10	14%	1/2%
9	20%	2%
8	25%	6%
7	21%	12%

1	2	3
6	12%	19%
5	5%	21%
4	2%	19%
3	1%	12%
2)	6%
1) negligible	2%
0)	½%

Discussion

The following observations may be made on these results :

1. The last column of Table II shows the effect of sampling errors. The candidates are *ex hypothesi* equally prepared, but only 21% receive their 'true' mark. The spread is very large, over 20% of the candidates having 7 marks or more, and another 20% 3 marks or less. There is no means of knowing which candidates are above or below their true score, and the examination is thus an extremely unreliable measure of the candidate's knowledge.

2. We may ask whether the provision of choice in a question paper compensates for the random element in the selection of questions. The table shows that this is not so. With choice, the central mark is now 8 (the true mean being 7.75). 20% of the candidates are still 2 marks or more below the central mark, and the marks are practically as widely spread as before.

The average performance, however, is nearly 3 marks better on the 9-question paper than on the 5-question paper. The performance of the candidates is therefore very seriously over-estimated by the examination. Choice is thus not a remedy for chance.

3. By assumption 7 the maximum possible mark is 15, and 5 may therefore be taken as the pass mark. On this basis only 3% of the

candidates fail on the 9-question paper. The proportion in conventional examinations is very much higher than this, which suggests that assumption 3 is too generous. Students doubtless get through these examinations with an even smaller proportion of total knowledge than is contemplated in this example.

4. In factual examinations, for which assumptions 1 and 2 are satisfied, choice should be abolished, since it distorts the estimate of the student's performance without improving its reliability. Reliability can be improved by increasing the number of questions and/or increasing the number of examinations.

5. In many examinations we are not testing for the range of factual knowledge. We may be testing a candidate's power of expression, or his ability to discuss evidence. The considerations of the present paper do not necessarily apply in such cases.

APPENDIX XIV

Evaluation of Essay-Type Answers

(Educational Testing Service, Princeton)

Readers were advised not to read a single paper more than once in order to arrive at separate ratings or percentages on each of the three criteria. It was felt that, with some practice, they could judge each paper as a whole, keeping in mind the relative weight of each of the major characteristics listed. Whether or not an individual reader made scratch-pad notes of a rating on each quality and summed them or do a like sum "in his head", or arrived at the same conclusion by intuition and broad generalization was considered a matter of personal preference.

Three comparison answers are first selected from among scripts valued by each of the moderators or examiners, as specimen valuations; the three scripts commonly agreed upon by them as representing the high, middle and low levels. Copies of these are made available to each examiner. The examiners can then study these three carefully, fix them in mind and refer to them with a minimum of effort. More distinctions will not be helpful since samples are clearly distinguishable at these levels. Supposing comparison essays are available at the levels 6, 4 and 2, "a score of 7 is awarded to a paper clearly superior to comparison essay 6, and a score of 1 is awarded to a paper clearly poorer than comparison essay 2. Scores of 5 mean that papers are not quite up to the mark of a six, but distinctly better than the sample of rating 4. Scores of 3 are awarded to papers that do not measure up to a 4, but that can be distinguished as better than a 2 sample."

By 'paper' in the above is meant an answer to a question, for it is in that context that the author is writing of a 'paper'. Once the

valuation of three samples of the answer to each question is agreed upon by the examiners concerned, they should start scoring the other scripts by comparing them with the three samples. This they should do "without regard to who wrote the essays or to his grade, or to his school." It would also be desirable to follow a specific procedure while scoring the scripts. Papers should be sorted into seven piles on the basis of their quality, as each paper gets valued, — corresponding to the rating 1 (low) to 7 (high) 4 is the middle level. "Ratings (or marks) should not be written on papers until the papers for a complete group have been sorted into piles." All that is stated in the preceding paragraphs is written with reference to a paper in which there is a single essay question. But it can be followed even with regard to papers which contain answers to more than one essay type question. All that has to be done is to place the paper at one of the seven levels on the basis of the quality of all the essay answers taken together, the quality of each answer being determined individually with reference to the three specimen answers that we have before us regarding that question.

Some papers, which are hard to assess on a first reading, will have to be reread. Such difficult papers should be marked at the time of first reading "with a rating that cannot be seen upon re-reading, and be put in a special 'eighth pile. After an interval of sufficient length to erase a specific recollection of the details, such papers can profitably be rescored, and the new score compared with the concealed original". If the two scores are two ratings apart (e.g., 5 and 3), the reader can assign the mid-rating (e.g. 4) to the paper. Unfortunately, if the paired scores are only one rating apart, it is impossible to "split the difference" since scores more definitive than the seven of the rating scheme cannot be interpreted.

The individual teacher scoring papers for students in his own classes should take precautions against "halo effect". This effect is the product of a teacher's expectation of performance — what he thinks each pupil can do, and should do in the circumstances. It is, of course, based upon a teacher's knowledge of past performance plus his judgment of ability, interest, and other factors. Any reading scheme, therefore, which identifies papers with their writers

raises the "halo" problem. The only answer is to recognize that factors other than the essay itself may influence grading and try consciously to rule out as many of these factors as possible.

The Essay tests should be graded on the basis of what the student says and how he says it. Under testing conditions, a teacher's normal insistence upon legibility and neatness cannot apply. Hence, teacher's normal insistence upon legibility and neatness cannot apply. Hence, teachers in the habit of "marking off" for sloppy work or giving a bonus for neatness are likely to fall into scoring error. But this weakness of a paper is an extraneous characteristic that should not be allowed to affect scores.

Other matters of personal bias should be examined with care and eliminated to the fullest extent practicable. These are such things as a desire to plow under all split infinitives with their splitters, or an especial aversion to the allergy to a half dozen specific clichés or to illiterate spelling. It is common for skilled readers to say that they must guard against a "pet peeve" of this sort or that. If the norms for the test are to be meaningful, the essay must be evaluated as a whole, and no one element should be decisive.

The above procedures will do for the single teacher who gives the test to his own students and must score the results himself. Greater reliability of scoring can be achieved for public examinations where some cooperation in scoring will have to be permitted. That is to say, there will be greater reliability if there are two examiners appointed for going through the same scripts and valuing them independently. The test will have been set at the same time and they will have to work together on the results.

Under a set-up of this kind, the following additional procedure can be observed:

When possible, two readings should be given each paper, with neither person knowing what the other gives as a score.

As much anonymity as possible should be built into the procedure. The papers should not be identified in terms of which

teacher's classes each comes from. (Even the work and reputation of teachers is subject to the "halo effect" mentioned above, and identification of classes from which papers come, and hence of the teacher will affect.

It has been found that more than two readings will not add a great deal to the reliability of the scores so little that a third reading is not justified except in especially hard-to-score cases.

The resolution of differences presents a problem in a formal double reading just as it does when the single teacher gets two different results. If scores are two ratings apart, the included score should be awarded; *i.e.* scores of 7 and 5 automatically are recorded as a 6, and scores of 4 and 2 become a 3. The problem is that most scores will be *one* rating apart, and cannot be split because on split scores can be interpreted. In these circumstances a reading by a third person is probably more satisfactory than the time-consuming process of review by both original readers.

