

INTERNATIONAL INSTITUTE EXAMINATIONS ENQUIRY

AN
EXAMINATION
OF
EXAMINATIONS

Being a Summary of Investigations on the Comparison of Marks allotted to Examination Scripts by Independent Examiners and Boards of Examiners, together with a Section on a Viva Voce Examination

BY

SIR PHILIP HARTOG, K.B.E., C.I.E.

AND

E. C. RHODES, D.Sc.,

READER IN STATISTICS IN THE UNIVERSITY OF LONDON

SECOND EDITION

Fourth Impression

MACMILLAN AND CO., LIMITED
ST. MARTIN'S STREET, LONDON

1936

Price One Shilling (By Post 1s. 2d.)

AN EXAMINATION OF EXAMINATIONS

INTERNATIONAL INSTITUTE EXAMINATIONS ENQUIRY

Members of the Committee:

SIR MICHAEL SADLER, K.C.S.I.,
C.B., LL.D. (*Chairman*)

Sometime Master of University
College, Oxford, and Vice-
Chancellor of the University of
Leeds.

P. B. BALLARD, M.A., D.Lit.

Sometime Inspector in the Educa-
tion Department of the London
County Council.

C. DELISLE BURNS, M.A., D.Lit.

Stevenson Lecturer in Citizen-
ship in the University of
Glasgow.

CYRIL BURT, M.A., D.Sc.

Professor of Psychology in the
University of London.

H. R. HAMLEY, M.Sc., Ph.D.

Professor of Education in the
University of London.

SIR PHILIP HARTOG, K.B.E., C.I.E.,
LL.D. (*Director*)

Sometime Vice-Chancellor of the
University of Dacca and Chair-
man of the Auxiliary Com-
mittee on Education of the
Indian Statutory Commission.

SIR PERCY NUNN, D.Sc., Litt.D.

Professor of Education in the
University of London.

C. SPEARMAN, LL.D., F.R.S.

Emeritus Professor of Psychology
in the University of London.

GODFREY H. THOMSON, D.Sc.

Professor of Education in the
University of Edinburgh.

F. CLARKE, M.A.

Professor-elect of Education in
the University of London.

Other publications of the Inter-
national Institute Examinations
Enquiry Committee:—

An English Bibliography of Ex-
aminations (1900-1932) by Mary
C. Champneys, with a Foreword
by Sir Michael Sadler and Sir
Philip Hartog (pp. xxiv, 141),
1934. Price 5/-.

Essays on Examinations, by Sir
Michael Sadler, A. Abbott,
P. B. Ballard, Cyril Burt,
C. Delisle Burns, Sir Philip
Hartog, C. Spearman and
S. D. Stirk (pp. xii, 166).
Price 5/-.

The Marks of Examiners, by Sir
Philip Hartog and Dr. E. C.
Rhodes, with a Memorandum by
Professor Cyril Burt. Price 8/6.

To be published shortly:—

A Conspectus of Examinations
conducted in Great Britain and
Northern Ireland.

INTERNATIONAL INSTITUTE EXAMINATIONS ENQUIRY

AN
EXAMINATION
OF
EXAMINATIONS

Being a Summary of Investigations on the Comparison of Marks allotted to Examination Scripts by Independent Examiners and Boards of Examiners, together with a Section on a Viva Voce Examination

BY

SIR PHILIP HARTOG, K.B.E., C.I.E.

AND

E. C. RHODES, D.Sc.,

READER IN STATISTICS IN THE UNIVERSITY OF LONDON

SECOND EDITION

Fourth Impression

MACMILLAN AND CO., LIMITED
ST. MARTIN'S STREET, LONDON

1936

WATERLOW and SONS LIMITED,
LONDON and DUNSTABLE.

TABLE OF CONTENTS

	PAGE
PREFACE TO FIRST EDITION - - - - -	6
PREFACE TO SECOND EDITION - - - - -	11

PART I.—GENERAL.

INTRODUCTION - - - - -	12
SCHOOL CERTIFICATE HISTORY - - - - -	14
SCHOOL CERTIFICATE LATIN - - - - -	16
SCHOOL CERTIFICATE FRENCH - - - - -	17
SCHOOL CERTIFICATE CHEMISTRY - - - - -	18
SCHOOL CERTIFICATE ENGLISH - - - - -	19
SPECIAL PLACE EXAMINATION (I): ARITHMETIC AND ENGLISH - - - - -	22
SPECIAL PLACE EXAMINATION (II): ENGLISH ESSAY - - - - -	26
COLLEGE ENTRANCE SCHOLARSHIP: ENGLISH ESSAY - - - - -	30
UNIVERSITY MATHEMATICAL HONOURS - - - - -	32
UNIVERSITY HISTORY HONOURS - - - - -	35
VIVA VOCE (INTERVIEW) EXAMINATION - - - - -	35

PART II.—DIFFERENCES OF STANDARD AND RANDOM

VARIATIONS OF DIFFERENT EXAMINERS - 42

SCHOOL CERTIFICATE HISTORY - - - - -	45
SCHOOL CERTIFICATE LATIN - - - - -	47
SCHOOL CERTIFICATE FRENCH - - - - -	48
SCHOOL CERTIFICATE CHEMISTRY - - - - -	51
SCHOOL CERTIFICATE ENGLISH - - - - -	52
SPECIAL PLACE EXAMINATION (II): ENGLISH ESSAY - - - - -	54
COLLEGE ENTRANCE SCHOLARSHIP: ENGLISH ESSAY - - - - -	54
UNIVERSITY MATHEMATICAL HONOURS - - - - -	55
UNIVERSITY HISTORY HONOURS - - - - -	56
SUMMARY OF THE FOREGOING SECTIONS - - - - -	56
METHOD OF CALCULATING IDEAL MARKS - - - - -	57

APPENDICES.

I.—UNIVERSITY HISTORY HONOURS (DETAILS OF INVESTIGATION)	59
II.—BRIEF SUMMARY OF THE WORK OF THE FRENCH INTERNATIONAL INSTITUTE EXAMINATIONS ENQUIRY - - - - -	78

PREFACE TO THE FIRST EDITION

1. No element in the structure of our national education occupies at the present moment more public attention than our system of examinations. It guards the gates that lead from elementary education to intermediate and secondary education, from secondary education to the Universities, the professions, and many business careers, from the elementary and middle stages of professional education to professional life.

2. Quite apart from the safeguards imposed by Acts of Parliament and Government authorities, a whole congeries of examinations has sprung up in the last century, created by private and public bodies¹. Examinations have become a familiar topic in our newspapers and in our homes. The examination system has grown to be an important element, not only in our education, but in the whole social system of our country; and the interest of many other countries in this matter is not less than our own.

3. The investigations on examinations of which this pamphlet is a summary are the outcome of an International Conference on Examinations held in May, 1931, at Eastbourne, under the auspices of the Carnegie Corporation, the Carnegie Foundation, and the International Institute of Teachers College, Columbia University. The countries represented at the Conference were (in alphabetical order) England, France, Germany, Scotland, Switzerland, and the United States². As a result of that

¹ In a *Conspectus* in preparation by the Committee there will appear between 150 and 200 names of such bodies, exclusive of Universities and Local Education Authorities.

² The Report of the Eastbourne Conference on Examinations, edited by Professor Paul Monroe, Director of the International Institute, was published by the Bureau of Publications, Teachers College, Columbia University, New York City, in 1931.

The representatives from the United States at the Conference were as follows:—

Dr. C. H. Judd, Dean of the School of Education, University of Chicago.

Dr. Frederick P. Keppel, President of the Carnegie Corporation, New York City.

Dr. Paul Monroe, Director of the International Institute, Teachers College, Columbia University.

Conference committees were set up in all the European countries above-named. Each of these committees received a grant for three years from the Carnegie Corporation through the International Institute, and each of them reported independently to a second International Conference held in June, 1935, at Folkestone, under the same auspices as the Conference held at Eastbourne. The Committees have done their work on independent lines and have reported separately. This pamphlet is substantially identical with the report presented by the English Committee to the Folkestone Conference, and it is published in its present form in accordance with a wish expressed at that Conference.

4. The English Committee consisted of the following: Sir Michael Sadler, K.C.S.I. (Chairman), Dr. P. B. Ballard, Dr. C. Delisle Burns, Professor Cyril Burt, Sir Philip Hartog, K.B.E. (Director), Professor Sir Percy Nunn, Professor C. Spearman, F.R.S., and Professor Graham Wallas. The Committee suffered a great loss in 1932 by the death of Professor Graham Wallas, who was replaced by Professor Godfrey Thomson, a member of the Scottish Committee. Professor H. R. Hamley and Professor C. W. Valentine joined the English Committee in the present year³. The address of the English Committee is 1, Plowden Buildings, Temple, London, E.C.4.

Dr. Henry Suzzallo, President of the Carnegie Foundation, New York City.

Dr. Edward L. Thorndike, Professor of Education, Teachers College, Columbia University.

* The membership of the other Committees is shown below:—

FRANCE—

M. A. Desclos, Directeur-adjoint de l'Office National des Universités et Ecoles Françaises (*President*).

M. Barrier, Adjoint au Directeur de l'Enseignement Primaire.

M. Bouglé, Directeur-adjoint de l'Ecole Normale Supérieure.

M. Gastinel, Inspecteur Général de l'Instruction Publique.

M. Laugier, Maître de Conférences à la Faculté des Sciences de Paris.

M. Luc, Directeur-adjoint de l'Enseignement Technique.

The original Committee included:

M. Charles Maurain, Doyen de la Faculté des Sciences de l'Université de Paris (who resigned on account of the pressure of other duties).

M. Cope, Président du Syndicat National des Professeurs des Lycées de Garçons et de l'Enseignement Secondaire Féminin (since deceased).

GERMANY—

Professor Erich Hylla, Ministerialrat im Ministerium für Kunst,

5. The Committee engaged Dr. E. C. Rhodes, Reader in Statistics in the University of London, to act as their statistician.

6. Touching education and social life as they do on so many points, the problems of examinations are many and varied. The Committee have published an *English Bibliography of Examina-*

Wissenschaft, und Volksbildung in Preussen; Professor an der Pädagogischen Akademie, Halle.

Dr. Robert Ulich, Ministerialrat im Ministerium für Volksbildung in Sachsen.

The original Committee included also:

Professor Dr. Carl Becker, Minister a.D. für Kunst, Wissenschaft, und Volksbildung in Preussen; Professor an der Universität, Berlin (since deceased).

Dr. Otto Bobertag, University of Berlin (since deceased).

SCOTLAND—

William Boyd, M.A., B.Sc., D.Phil., Lecturer in Education, Glasgow University.

Shepherd Dawson, M.A., D.Sc., Lecturer in Psychology, Jordanhill Training College, Glasgow (since deceased).

Professor James Drever, M.A., D.Phil., Professor of Psychology, Edinburgh University.

Thomas Henderson, B.Sc., F.E.I.S., Hon. Secretary of the Scottish Council for Research in Education.

W. A. F. Hepburn, M.C., M.A., B.Ed., Director of Education to the Ayrshire Education Committee.

Professor W. W. McClelland, M.A., B.Sc., B.Ed., Professor of Education, St. Andrews University.

J. Mackie, M.A., D.Sc., F.R.S.E., Head Master, Leith Academy.

Robert R. Rusk, M.A., B.A., Ph.D., Lecturer in Education, Jordanhill Training College, Glasgow; Director to the Scottish Council for Research in Education.

J. C. Smith, C.B.E., M.A., D.Litt., formerly Senior Chief Inspector of Schools, Scottish Education Department.

Professor Godfrey H. Thomson, Ph.D., D.Sc., Professor of Education, Edinburgh University.

SWITZERLAND—

M. Pierre Bovet, Professeur à l'Université de Genève; Directeur de l'Institut Universitaire des Sciences de l'Education, Genève.

Dr. Brenner, Directeur du Lehrerseminar, Bâle.

M. Edouard Claparède, Professeur de Psychologie à l'Université de Genève; Directeur de l'Institut Jean-Jacques Rousseau.

M. Robert Dottrens, Directeur d'Ecoles, Troinex, Genève (Dr. Soc.).

Dr. Charles Junod.

M. Albert Malche, Conseiller aux Etats; Professeur à l'Université de Genève.

M. Jean Piaget, Directeur du Bureau International d'Education.

tions (1900-32)⁴, which shows how much has been written on the subject in this country during the first third of the century. The Committee are also publishing a volume of *Essays on Examinations*, dealing with a number of aspects of the subject, which will appear soon after this pamphlet, and a *Conspectus of Examinations in Great Britain and Northern Ireland*, which will appear later. But the main work carried out for the Committee will be recorded in a volume entitled *The Marks of Examiners*, now in course of printing, of which the present pamphlet is a summary.

7. The object of the investigations to be described may be explained very simply. Professor F. Y. Edgeworth, many years ago, found that the marks allotted independently by twenty-eight different examiners to a piece of Latin prose varied from 45 to 100 per cent. In the United States, Messrs. Starch and Elliott, and, in France, M. Laugier and Mlle. Weinberg have found similar results, but no systematic comparison has hitherto been published of the marks allotted by a number of different examiners, all experienced and qualified for their task, to sets of scripts (answer-books) actually written at public examinations. Both the English and the French Committees have attacked this subject, and the present pamphlet gives a fairly extended summary of the English results and a brief one of the French. These results are similar in the two countries, and equally disquieting. It is clear that the part played by chance in the verdicts given at different examinations on which careers depend must often at the present moment be a great one. The Committee are well aware that the consideration of borderline cases by examination authorities does materially diminish the chances of a candidate being wrongly rejected; but it must be pointed out that candidates may be placed in error below the

Genève; Professeur extraordinaire à l'Université de Genève;
Co-directeur de l'Institut Jean-Jacques Rousseau.

Dr. W. Schohaus, Schweizerische Erziehungs Rundschau, Kreuzlingen, Thurgovie.

Dr. Ida Somazzi, Seminar, Berne.

Dr. Hans Stettbacher, Lehramtkurse, Universität, Zurich.

M. Teodoro Valentini, Professeur, Scuola Normale, Locarno, Tessin.

⁴ *An English Bibliography of Examinations (1900-1932)*, by Mary C. Champneys, with a Foreword by Sir Michael Sadler and Sir Philip Hartog (Macmillan & Co., Ltd.), 1934.

“borderline.” Again, it must be remembered in the interest of the public, to whom an examination certificate means a certificate of efficiency, that candidates may now by chance obtain such certificates when they should by rights be rejected.

8. Of all the results recorded by the English Committee perhaps the most disturbing are those recorded in the investigation on the marking of School Certificate History scripts. It was found that when fourteen experienced examiners re-marked independently fifteen scripts which had all received the same moderate mark from the examining authority by which they were furnished, these examiners, between them, allotted over forty different marks to the several scripts. It was found, further, that when these examiners re-marked once more the same scripts after intervals of from twelve to nineteen months, they changed their minds as to the verdict of Pass, Fail, and Credit in 92 cases out of the total of 210. Clearly a test of this kind cannot inspire confidence.

9. Our investigations show that the employment of boards of examiners instead of individual examiners, though it diminishes, does not remove the element of chance in examinations, and that boards, as well as individuals, may disagree in their verdicts. The element of chance in examinations still subsists to a dangerous degree in the subjects which have been investigated by the Committee.

10. The question may at once be asked : Should examinations be abolished ? If not, what remedies can be suggested ?

The Committee are clearly opposed to the root and branch policy. They are of opinion that examinations as a test of efficiency are necessary. They are further of opinion that, in addition to those examinations which yield identical results when applied by different examiners (e.g. “New Type” or “Objective” examinations), the traditional “essay” examination should be preserved. But they hold that it is as impracticable to recommend an *a priori* cure for the defects of the present examination system as it would be to recommend an *a priori* cure for a disease. It is only by careful and systematic experiment that methods of examination can be devised not liable to the distressing uncertainties of the present system. No doubt investigations like those recorded by our Committee, and administrative experiments in allowing teachers, in conjunction with Government or University inspectors, to “brand

their own herrings," would involve expenditure, but such expenditure and experiments would be justified in the public interest.

The Committee desire to acknowledge their deep obligation to the various examination authorities by whom they have been furnished with the scripts which formed the material for their investigations, or by whom they have been assisted in other ways, and to the examiners who marked the scripts or took part in the *circa voce* examination. Without the cordial assistance both of examination authorities and of examiners, it would have been impossible for the Committee to carry out their investigations on the lines which they had planned.

In conclusion, the Committee wish to express their warm appreciation of the generosity and initiative of the Carnegie Corporation, the Carnegie Foundation, and the International Institute of Teachers College, Columbia University, to which this Committee and the parallel Committees in other countries owe their existence.

PREFACE TO THE SECOND EDITION

The first edition of this pamphlet appeared in December, 1935. A few slips were corrected in the second impression, which was issued shortly afterwards; and some further corrections of detail have been made in the present edition. These corrections do not in any way affect the conclusions of the Committee. While the pamphlet has been received with warm approval by the general public, it has evoked certain criticisms, with some of which it is proposed to deal in *The Marks of Examiners*, now nearly ready for publication. To have dealt with the criticisms in this pamphlet would have involved an increase in both its size and price which was thought undesirable.

It should be added that Professor Valentine, who was elected a member of the Committee in July, 1935, resigned at the end of December in the same year, and that Professor F. Clarke, M.A., Advisor to the Overseas Students in the Institute of Education of the University of London and Director-elect of the Institute, who has been in close touch with the Committee for some time, became a member early in 1936.

April, 1936.

APPENDIX I

UNIVERSITY HISTORY HONOURS (DETAILS OF INVESTIGATION)

1. *Character of Examination Papers.*—The examination papers were four in number, all forming part of a University History Honours Examination. The subjects of the papers were as follows :—

Paper I. Ancient and Mediæval History.

Paper II. Mediæval and Modern History.

Paper III. An Essay-paper with a choice from a number of subjects.

Paper IV. Political Thought (Prescribed Books).

In Papers I, II, and IV, candidates were requested not to attempt more than four questions out of a considerable number. The time allowed for each paper was three hours.

2. *Procedure.*—The University concerned furnished us with all the scripts available in the subjects enumerated above from a recent Honours examination.¹ Unfortunately 3 scripts (which happened to be among the best) had been accidentally destroyed. The total number of scripts available was 18 for Paper I, 17 for Paper II, 18 for Paper III, and 16 for Paper IV.

¹The examination included a number of other papers, but it was thought that the field covered by these was sufficient for the purpose of the investigation.

The following 17 examiners took part in the marking of the scripts :—

PROFESSOR J. B. BLACK, M.A., Burnett-Fletcher Professor of History in the University of Aberdeen.

PROFESSOR A. BROWNING, M.A., D.Litt., Professor of History in the University of Glasgow.

MR. NOEL DENHOLM-YOUNG, M.A., Fellow of Magdalen College, Oxford.

PROFESSOR A. H. DODD, M.A., Professor of History in the University of Wales.

MR. D. L. KEIR, M.A., Fellow of University College and University Lecturer in English Constitutional History, Oxford.

MR. R. B. McCALLUM, M.A., Fellow and Lecturer in Modern History, Pembroke College, Oxford.

PROFESSOR J. L. MORISON, M.A., D.Litt., Professor of Modern History, Armstrong College, University of Durham.

PROFESSOR R. B. MOWAT, M.A., Professor of History in the University of Bristol.

MR. J. N. L. MYRES, M.A., Student and Tutor of Christ Church, Oxford.

MR. E. J. PASSANT, M.A., Fellow of Sidney Sussex College, Cambridge.

MISS I. G. POWELL, M.A., Lecturer in History at the Royal Holloway College, University of London.

PROFESSOR EILEEN POWER, M.A., D.Lit., Professor of Economic History in the University of London.

PROFESSOR F. M. POWICKE, Litt.D., F.B.A., Regius Professor of Modern History in the University of Oxford.

MR. G. H. STEVENSON, M.A., Fellow of University College and University Lecturer in Ancient History, Oxford.

MR. C. G. STONE, M.A., Balliol College, Oxford.

PROFESSOR A. F. BASIL WILLIAMS, O.B.E., M.A., F.B.A.,
Professor of History in the University of Edinburgh.

PROFESSOR C. H. WILLIAMS, M.A., Professor of History in
the University of London.

The examiners are designated A, B, C, . . . R, in what follows, but this designation does not correspond with the alphabetical order of the names.

3. The scripts of Paper I were marked by 5 examiners; the scripts of each of the other papers by 10 examiners. The only reason for having the scripts of Paper I marked by fewer examiners was the difficulty in getting examiners to cover the two periods with which it dealt.

As in other investigations, no indication of origin or of the original marking appeared on the scripts, or was communicated to the examiners.

Each examiner marked each individual question separately and gave a final mark for each script as a whole.

4. The following "literal" system of marking, including 24 grades ranging from δ to $\alpha+$, was, after consultation with an eminent historian, submitted to and approved by the great majority of examiners before the work began. It was communicated as approved to one or two examiners who came into the investigation subsequently.

TABLE 1

<i>Literal Mark</i>	<i>No. of Grade</i>	<i>Literal Mark</i>	<i>No. of Grade</i>	<i>Literal Mark</i>	<i>No. of Grade</i>
$\alpha+$	(24)	$\beta++$	(15)	$\beta\gamma$	(6)
$\alpha?+$	(23)	$\beta+?+$	(14)	$\gamma\beta$	(5)
α	(22)	$\beta+$	(13)	$\gamma+$	(4)
$\alpha?-$	(21)	$\beta?+$	(12)	γ	(3)
$\alpha-$	(20)	β	(11)	$\gamma-$	(2)
$\alpha-?-$	(19)	$\beta?-$	(10)	δ	(1)
$\alpha=$	(18)	$\beta-$	(9)		
$\alpha\beta$	(17)	$\beta-?-$	(8)		
$\beta\alpha$	(16)	$\beta=$	(7)		

5. It may be well to say a word here on the use of a literal system of this kind as compared with the numerical systems employed in our other investigations. The literal system is generally used at Oxford; there is a considerable variety of usage in other Universities.

6. There seems to be a fundamental difference, at any rate at the first blush, between the two systems. The literal system indicates only an order in classification, not ratios of proficiency. With that system, there can be no question of adding up marks for individual questions in order to obtain a percentage of a total maximum. It would appear that the literal mark indicates in the examiner's mind a certain "quality." The question of "quantity" probably enters into his estimate only in a subordinate degree.

With the numerical system, on the other hand, the marks for individual questions are added up to furnish a total, a procedure which is convenient, though it is based on hypotheses which it is not perhaps easy to analyse and justify. But any attempt to add together the symbols indicating "classes" or "grades" would seem *a priori* unjustifiable and would be rejected by many who use literal marks.

7. Both systems have their conveniences. It is for the sake of readers who are unaccustomed to literal marking, and to enable them to estimate by what number of grades (or subordinate classes) any two examiners differ, that we have attributed the numbers 1 to 24 to the successive grades, δ to $\alpha+$, and that, side by side with the literal tables, we have inserted numerical tables on this basis. But, for the reasons stated above, the numbers indicating grades must not be regarded as numerical marks. They are ordinal numbers, not cardinal.

8. Readers accustomed to numerical marking may further wish to have some means of comparison between the two systems. A rough and ready form of translation from one into the other would be to suppose that each of the 24 literal symbols corresponds to a multiple of four marks, and the highest, $\alpha+$, to 96. Only an experimental investigation could afford any real basis for such a translation. But it is certain that such a difference as that of 18 grades, the maximum difference between the awards of two different examiners to the same script in this investigation, much more nearly approaches a difference of 72 in numerical

marking, with 96 (or 100) as a maximum mark, than a difference of 18, which a superficial glance might suggest.

9. An index of the examiners who marked the various papers is given in the Table below :—

TABLE 2

<i>Examiner</i>	<i>Paper</i>			
	I	II	III	IV
A	-	*	*	*
B	-	*	*	*
C	-	*	-	-
D	*	-	-	-
E	-	-	-	*
F	-	*	*	*
G	-	-	-	*
H	-	*	*	*
J	-	*	*	*
K	*	*	*	-
L	-	*	*	*
M	-	-	-	*
N	-	*	*	-
O	*	-	-	-
P	*	-	-	-
Q	*	-	*	*
R	-	*	*	-

The papers marked by each examiner are indicated by an asterisk in the row corresponding to the letter by which he is designated. Thus Examiner B marked Papers II, III and IV.

10. In Tables 3, 4, 4a, 5, 5a, 6 and 6a are set out the literal marks assigned by the examiners to the scripts of each candidate, and the numerical representation of the corresponding grades according to the convention explained in paras. 7 and 8 above.

TABLE 3.

Paper I.

Marks allotted						Numerical representation of the the marks in ordered grades						Range in grades	Range in grades neglecting Q's results
Examiner	D	K	O	P	Q	D	K	O	P	Q			
Cand. No.	1	$\beta++$	$\beta?+$	β	$\beta=$	$\gamma\beta$	15	12	11	7	5	10	8
	2	$\alpha\beta$	$\beta?-$	$\beta+$	$\beta?+$	β	17	10	13	12	11	7	7
	3	$\beta?-$	$\beta+$	$\beta?-$	$\beta-$	$\gamma\beta$	10	13	10	9	5	8	4
	4	$\beta-$	$\beta++$	β	$\beta?-$	γ	9	15	11	10	3	12	6
	5	$\beta+$	$\beta?-$	$\beta-$	$\gamma+$	$\gamma+$	13	10	9	4	4	9	9
	6	$\alpha\beta$	$\beta+?+$	$\alpha-$	$\beta+$	$\gamma\beta$	17	14	20	13	5	15	7
	7	$\beta-$	$\beta?+$	$\beta?+$	$\gamma\beta$	$\beta\gamma$	9	12	12	5	6	7	7
	8	$\beta?+$	$\beta+$	β	$\gamma+$	$\gamma+$	12	13	11	4	4	9	9
	9	$\alpha-$	β	$\beta-$	$\beta+$	$\beta+$	20	11	9	13	13	11	11
	10	$\beta-$	$\beta+$	β	$\beta-$	$\beta?+$	9	13	11	9	12	4	4
	11	γ	$\gamma-$	$\gamma-$	$\gamma-$	$\gamma-$	3	3	2	2	2	1	1
	12	$\gamma\beta$	$\beta+?+$	β	$\beta-$	$\gamma\beta$	5	14	11	9	5	9	9
	13	$\alpha\beta$	$\alpha\beta$	α	$\gamma\beta$	$\beta\gamma$	17	17	22	5	6	17	17
	14	$\beta\alpha$	$\beta++$	$\beta++$	$\beta+$	$\beta?-$	16	15	15	13	10	6	3
	15	$\beta-$	$\gamma+$	β	γ	γ	9	4	11	3	3	8	8
	16	$\beta-$	$\beta+?+$	β	γ	$\beta+$	11	14	11	3	13	11	11
	17	$\beta=$	β	β	$\beta+$	$\gamma\beta$	7	11	11	13	5	8	6
	18	α	$\beta++$	$\alpha?-$	β	β	22	15	21	11	11	11	11
Median		β $\beta?+$	$\beta+$	β	$\beta-$	$\gamma\beta$	11-12	13	11	9	5	Average 9.1	Average 7.7

TABLE 4

Paper II

Marks allotted

Examiner	A	B	C	F	H	J	K	L	N	R
Cand. No. 1	$\beta+?+$	$\beta-$	$\alpha-?+$	$\beta-$	$\beta\alpha$	β	β	$\beta\gamma$	$\gamma\beta$	$\beta?-$
2	$\beta?+$	$\beta+$	$\beta\alpha$	$\alpha\beta$	$\beta++$	$\beta+$	$\beta\alpha$	$\beta-$	$\beta+?+$	$\beta-$
3	β	$\beta+?+$	β	$\beta-$	$\beta+$	$\beta-$	$\beta++$	$\beta=$	$\beta\gamma$	$\beta\alpha$
4	$\beta-$	$\alpha=$	$\beta\beta$	$\beta+?+$	$\beta+$	$\gamma+$	$\beta+?+$	$\beta-$	$\beta-$	$\beta+$
5	$\beta++$	$\beta++$	$\beta?-$	$\beta?+$	$\beta?+$	$\gamma+$	$\beta?-$	$\beta-$	$\gamma\beta$	$\beta?+$
6	$\beta++$	$\alpha-?+$	$\beta?+$	$\beta+?+$	$\beta+?+$	$\beta?+$	$\beta-$	$\beta+$	β	$\beta++$
7	$\beta\alpha$	$\alpha\beta$	$\beta+$	$\beta?+$	$\beta-$	γ	$\beta-$	$\beta+$	β	$\beta++$
8	β	$\alpha-$	β	$\beta?+$	$\beta-$	$\gamma+$	$\beta+$	$\beta-$	$\beta+?+$	$\beta++$
9	$\alpha=$	$\alpha\beta$	$\beta-$	$\beta\gamma$	$\beta\gamma$	β	$\beta+$	$\beta-$	$\beta+$	$\alpha\beta$
10	$\beta+$	$\beta+$	$\beta+?+$	$\beta\alpha$	$\gamma\beta$	$\beta-$	$\beta+?+$	$\beta=$	$\beta\beta$	$\beta++$
11	β	$\beta=$	$\beta-$	$\beta-?+$	$\gamma+$	$\beta-$	$\beta\gamma$	$\beta\gamma$	β	$\gamma\beta$
12	$\beta?+$	$\beta?+$	$\alpha\beta$	β	$\gamma+$	β	$\beta+?+$	$\beta-$	$\beta++$	$\beta\beta$
13	$\beta++$	$\beta+$	$\beta?+$	$\beta++$	$\beta?+$	$\gamma\beta$	$\beta++$	$\beta-$	$\beta++$	$\beta+?+$
14	$\beta+?+$	$\alpha-$	$\alpha=$	$\beta\alpha$	$\gamma\beta$	$\beta++$	$\beta\alpha$	$\beta+$	$\beta\alpha$	$\beta+$
15	$\beta?+$	$\gamma+$	$\beta?+$	$\beta-$	$\gamma?+$	$\beta=$	$\gamma+$	$\beta+$	$\beta=$	$\beta?+$
17	$\beta?+$	$\beta+$	β	$\beta+$	$\beta?+$	$\beta=$	$\beta?+$	$\beta=$	β	$\beta?+$
18	$\beta++$	$\beta++$	$\beta?+$	$\beta\alpha$	$\beta?+$	$\beta+$	$\beta+$	$\beta+?+$	$\beta\alpha$	$\alpha-$
Median	$\beta+$	$\beta+?+$	$\beta?+$	$\beta?+$	$\beta?+$	$\beta-$	$\beta+$	$\beta?+$	β	$\beta+$

AN EXAMINATION OF EXAMINATIONS

TABLE 1.4A

Paper II

66

Numerical representation of the marks in ordered grades

Examiner	A	B	C	F	H	J	K	L	N	R	Range in grades
Cand. No. 1	14	9	19	9	16	11	11	6	5	10	14
2	12	13	16	17	15	13	16	9	14	9	8
3	11	14	11	9	13	9	15	7	6	16	10
4	9	18	17	14	13	4	14	10	11	13	14
5	15	15	10	12	12	4	10	10	5	12	11
6	15	19	12	14	14	12	9	13	11	15	10
7	16	17	13	10	6	9	3	12	11	12	14
8	11	20	11	10	9	4	13	9	14	15	16
9	18	17	9	6	6	11	13	10	13	17	12
10	13	13	14	16	5	9	14	7	5	15	11
11	1	7	9	8	4	9	6	6	2	5	8
12	12	12	17	11	4	11	14	10	15	17	13
13	15	13	10	15	12	5	15	9	12	14	10
14	14	20	18	16	5	15	16	13	16	13	15
15	10	4	10	9	3 $\frac{1}{2}$	7	4	12	7	10	8 $\frac{1}{2}$
17	12	13	11	13	10	7	12	7	11	12	6
18	15	15	12	16	12	13	13	14	16	20	8
Median	13	14	12	12	10	9	13	10	11	13	Average 11.1

AN EXAMINATION OF EXAMINATIONS

TABLE 5

Paper III

Marks allotted.

Examiner	A	B	F	H	J	K	L	N	Q	R
Cand. No. 1	$\beta\alpha$	$\beta?+$	β	$\alpha=$	$\beta+$	$\beta+$	$\beta?+$	β	$\gamma\beta$	$\beta+$
2	$\beta-$	$\beta+$	$\alpha=$	$\beta+$	$\beta?+$	β	$\beta?+$	$\beta-$	$\beta+?+$	$\beta+$
3	$\beta?+$	$\beta++$	β	$\beta++$	β	$\beta+?+$	$\beta+$	$\beta\gamma$	$\beta\gamma$	α
4	$\beta?-$	β	β	β	$\beta?+$	β	$\beta?-$	$\beta\gamma$	γ	$\beta?+$
5	$\beta\alpha$	$\gamma+$	$\beta-$	$\beta-$	$\beta+$	$\beta+$	$\beta?-$	$\gamma\beta$	$\gamma+$	$\beta-$
6	$\beta++$	$\alpha-$	$\beta\alpha$	$\beta+$	$\beta++$	$\beta?+$	$\beta+?+$	$\beta?+$	$\gamma\beta$	$\alpha\beta$
7	β	β	$\beta-$	β	$\beta-$	$\beta=$	β	$\beta=$	$\beta\gamma$	$\beta+$
8	β	$\alpha=$	β	β	$\beta+$	$\beta-$	$\beta+$	$\beta=$	$\beta\gamma$	$\beta+$
9	α	$\gamma+$	$\beta+$	$\gamma\beta$	$\alpha-$	$\beta+?+$	$\beta+$	$\beta+$	$\alpha\beta$	$\alpha-$
10	$\beta?+$	$\beta+$	$\beta++$	$\beta-$	$\gamma+$	$\beta++$	$\beta?-$	β	$\beta++$	$\beta?+$
11	δ	γ	$\beta=$	β	$\gamma-$	$\beta\gamma$	$\gamma\beta$	β	$\gamma-$	$\gamma+$
12	$\beta?+$	$\gamma+$	$\beta+$	β	β	$\beta?-$	$\beta\gamma$	$\beta=$	$\gamma+$	β
13	$\beta+$	$\beta-$	$\alpha-$	$\beta+$	$\beta+$	$\beta\alpha$	$\beta?-$	$\beta+$	$\beta-$	$\beta?+$
14	$\beta+?+$	$\beta\alpha$	$\alpha\beta$	$\gamma+$	$\gamma+$	$\beta+?+$	$\beta?+$	β	$\beta+$	$\beta\alpha$
15	β	$\beta++$	$\beta-$	$\gamma\beta$	β	$\gamma+$	$\beta=$	β	$\gamma+$	$\beta+?+$
16	$\beta+$	$\alpha-$	$\alpha\beta$	$\beta\alpha$	$\beta\alpha$	$\beta?+$	$\beta\alpha$	$\beta+?+$	$\alpha\beta$	$\beta-$
17	$\beta?+$	$\gamma+$	$\beta=$	$\beta?+$	γ	γ	$\beta\gamma$	γ	γ	$\beta?-$
18	β	$\beta++$	$\alpha=$	$\beta++$	$\beta++$	β	$\beta\alpha$	$\beta+$	$\beta\alpha$	$\alpha?+$
Median	$\beta?+$	$\beta+ \}$ $\beta?+ \}$	$\beta?+$	β	$\beta?+$	$\beta \}$ $\beta?+ \}$	$\beta \}$ $\beta?+ \}$	$\beta?-$	$\beta\gamma$	$\beta+$

AN EXAMINATION OF EXAMINATIONS

67

TABLE 5A

Paper III

68

Numerical representation of the marks in ordered grades.

<i>Examiner</i>	A	B	F	H	J	K	L	N	Q	R	<i>Range in grades</i>	<i>Range in grades neglecting Q's results</i>
Cand. No. 1	16	12	11	18	13	13	12	11	5	13	13	7
2	9	13	18	13	12	11	12	9	14	13	9	9
3	12	15	11	15	11	14	13	6	6	22	16	16
4	10	11	11	11	12	11	10	6	3	12	9	6
5	16	4	9	9	13	13	10	5	4	9	12	12
6	15	20	16	13	15	12	14	12	5	17	15	8
7	11	11	9	11	9	7	11	7	6	13	7	6
8	11	18	11	11	13	9	13	7	8	13	12	11
9	22	4	13	5	20	14	13	13	17	20	18	18
10	12	13	15	9	4	15	10	11	15	12	11	11
11	1	3	7	11	2	6	5	1	2	4	10	10
12	12	4	13	11	11	10	6	7	4	11	9	9
13	13	9	20	13	13	16	10	13	9	12	11	11
14	14	16	17	4	4	14	12	11	13	16	13	13
15	11	15	9	5	11	4	7	11	4	14	11	11
16	13	20	17	16	16	12	16	14	17	15½	8	8
17	12	4	7	12	3	3	6	3	3	10	9	9
18	11	15	18	15	15	11	16	13	16	23	12	12
Median	12	12-13	12	11	12	11-12	11-12	10	6	13	Average 11-4	Average 10-4

AN EXAMINATION OF EXAMINATIONS

TABLE 6

Paper IV

Marks allotted

Examiner	A	B	E	F	G	H	J	L	M	Q
Cand. No. 1	β	$\beta^?+$	$\beta-$	$\beta^?-$	$\beta^?-$	$\beta^?+$	$\beta^?++$	$\beta^?-$	$\beta^?++$	$\beta\gamma$
2	$\beta^?-$	$\beta-$	$\beta^?+$	$\beta\alpha$	$\beta+$	$\beta+$	$\alpha=$	$\beta^?+$	$\beta+$	$\beta^?+$
3	$\beta+$	$\beta+$	$\beta+$	β	$\beta^?++$	$\beta-$	$\beta^?+$	$\beta^?+$	$\beta^?-$	$\beta-$
4	$\beta^?++$	β	$\beta\alpha$	$\beta^?+$	$\beta-?-$	β	$\beta+$	$\beta\gamma$	$\beta+$	$\beta\gamma$
5	β	$\beta^?-$	$\beta^?-$	$\beta=$	$\beta-$	$\gamma+$	$\gamma+$	$\gamma\beta$	$\gamma\beta$	$\gamma+$
6	$\beta\alpha$	$\beta+$	$\beta^?-$	$\beta^?-$	$\alpha=$	$\beta\gamma$	$\beta^?+$	$\beta^?+$	$\beta^?++$	$\gamma\beta$
7	$\beta\gamma$	$\gamma+$	$\beta=$	$\beta^?+$	$\beta^?+$	$\beta\gamma$	$\gamma+$	$\beta=$	$\beta=$	γ
8	$\beta\gamma$	$\alpha-$	$\beta+$	$\beta+$	$\beta=$	$\beta^?+$	$\gamma+$	$\beta^?+$	$\beta+$	$\beta\gamma$
9	α	$\beta+$	$\beta\alpha$	$\beta^?-$	$\alpha\beta$	$\beta^?+$	$\beta+$	$\beta+$	$\beta\alpha$	$\beta^?++$
10	$\beta\alpha$	$\beta^?+$	$\beta^?+$	$\alpha\beta$	$\beta^?-$	$\beta+$	$\beta^?+$	$\beta-$	$\beta+$	$\beta^?++$
11	δ	γ	$\gamma+$	$\beta^?-$	$\beta-$	γ	$\gamma+$	$\gamma\beta$	$\beta=$	$\beta\gamma$
12	$\beta^?++$	β	$\beta^?+$	$\beta^?+$	$\beta^?+$	$\beta\gamma$	$\beta+$	$\beta\gamma$	β	$\gamma+$
13	$\beta\alpha$	$\alpha-$	$\alpha-?-$	$\alpha=$	$\beta^?+$	$\beta^?+$	$\beta^?++$	$\beta=$	$\alpha\beta$	$\beta+$
15	$\beta^?++$	$\beta+$	$\alpha\beta$	$\beta^?+$	$\beta+$	$\beta+$	$\gamma+$	$\beta-$	$\beta^?+$	$\beta-$
17	β	β	$\beta-$	$\beta-?-$	$\beta+$	$\beta=$	$\beta+$	$\beta=$	$\beta+$	$\gamma+$
18	$\beta^?-$	$\beta^?+$	$\alpha\beta$	$\beta\alpha$	$\alpha\beta$	$\alpha-$	$\beta^?+$	$\beta^?++$	$\beta^?++$	$\alpha\beta$
Median	$\beta^?+$	$\beta^?+$	$\beta^?+$	β $\beta^?+$	$\beta^?+$	β $\beta^?+$	$\beta+$	$\beta-$	$\beta+$	$\beta\gamma$

AN EXAMINATION OF EXAMINATORS

TABLE 6A

Paper IV

Numerical representation of the marks in ordered grades

Examiner	A	B	E	F	G	H	J	L	M	Q	Range in grades	Range in grades neglecting Q's results
Cand. No. 1	11	12	9	10	10	12	14	10	14	6	8	5
2	10	9	12	16	13	13	18	12	13	12	9	9
3	13	13	13	11	14	9	15	12	10	9	6	6
4	14	11	16	12	8	11	13	6	13	6	10	10
5	11	10	10	7	9	4	4	5	5	4	7	7
6	16	13	10	10	18	6	15	12	14	5	13	12
7	6	4	7	10	12	6	4	7	11	3	9	8
8	6	20	13	13	7	12	4	12	13	6	16	16
9	22	13	16	10	17	12	13	13	16	14	12	12
10	16	12	12	17	10	13	12	9	13	14	8	8
11	1	3	4	10	9	3	4	5	7	6	9	9
12	14	11	12	12	12	6	13	6	11	4	10	8
13	16	20	19	18	12	15	14	7	17	13	13	13
15	14	13	17	12	13	13	4	9	15	9	13	13
17	11	11	9	8	13	7	13	7	13	4	9	6
18	10	15	17	16	17	20	12	14	14	17	10	10
Median	12	12	12	11-12	12	11-12	13	9	13	6	Average 10.1	Average 9.5

11. A glance at the Tables shows certain general features of interest. We have a closeness of marking between certain examiners and a wide difference between others, not attributable to chance, but showing real and probably irreconcilable differences of standard.

12. The examiners were asked to indicate what were their limits for a First, a Second, and a Third Class. Not all replied on the point. In the original scheme, a copy of which was furnished to each examiner (see para. 4), there was a gap between $\beta\alpha$ and $\beta++$, and between $\beta=$ and $\beta\gamma$, there being tacitly implied three classes. The following is a summary of information supplied by the examiners on the meaning of the symbols.

- A :— $\alpha\beta$ and $\beta\alpha$ borderline. So also $\beta\gamma$ and $\gamma\beta$. δ fails.
- B :—Nil.
- C :— $\alpha\beta$ is a first, $\beta\alpha$ a second, $\beta\gamma$ is a third class. δ is a failure. Rarely uses high α 's, or low marks, e.g. γ 's.
- D :—Does not use $\alpha+$ or $\alpha++$, perfection is α . $\beta\alpha$ or $\beta++$ is the best second class. He would have put $\beta\alpha$ at the top of the second group.
- E :— $\alpha\beta$ and $\beta\alpha$ are borderline marks, the former indicating a first class paper with either one poor answer or one persistent fault, the other a second class paper with one excellent answer or one very sound quality. Similarly with other borderline marks. Failures are $\gamma-$ and δ .
- F :— $\beta\alpha$ is top of second class. $\beta\gamma-$ is top of third. δ is failure.
- G :— $\beta\alpha$ is top of second, $\beta=$ is top of third class, $\alpha\beta$ and $\beta\alpha$ are borderline and $\beta-f-$ is borderline. $\gamma-$ and δ are failures.
- H :— $\alpha\beta$ and $\beta\alpha$ as in E. δ is failure.
- J :—First, second and third class as implied in the scheme sent out.
- K :—Nil.
- L :— $\alpha\beta$ minimum for first class. $\beta\alpha$ borderline. $\beta\gamma$ minimum for second. $\gamma\beta$ borderline. δ failure.
- M :— $\alpha\beta$ minimum for first class. $\beta\alpha$ borderline. $\beta\gamma$ and $\gamma\beta$ borderline. δ failure.
- N :— $\alpha\beta$ minimum for first class. $\beta\alpha$, second; third, $\beta\gamma$ to and including δ .
- O :—As in E with qualification "that value of borderline marks as means of judging is that, if several papers have to be assessed in the final result, the mixed or "border" marks have an additional significance, pointing to the need for inquiry. They suggest quality. Hence I should personally avoid them if only one paper was set on a subject."

P:— $\alpha\beta$ and $\beta\alpha$ borderline as E. So with the $\beta\gamma$ and $\gamma\beta$.

Q:— $\alpha\beta$ minimum for first class. $\beta\alpha$ highest second. So with others.

R:— $\alpha\beta$ and $\beta\alpha$ borderline. $\beta-$, $\beta-?-$, $\beta=$ borderline. $\beta\gamma$ highest third class. $\gamma-$ and δ fail.

13. The examiners are not in sufficient agreement on this point to use their remarks as a basis for classification. In actual practice it is well-known that the limits are not determined in any purely mechanical way, but are the subjects of discussion in connexion with all border line cases. The subject of the present investigation is not the actual award of First, Second and Third Classes at a History Honours examination, but the variation in the individual judgments which must serve as a basis for those awards.

Although we cannot use the terms First, Second and Third Class, we can distinguish between the number of α 's, β 's, γ 's, and δ 's and of borderlines.

Thus the lowest limit for a First Class most generally adopted is $\alpha\beta$; but some are willing to consider $\beta\alpha$, the next grade, as a borderline for a First.

There is much more variation in the opinions as to the lower limit of a Second Class:—

β is adopted by F,

$\beta=$, by C, H, J, and N.

$\beta\gamma$, by Q.

Some of the other examiners indicate that the borderline marks between second and third class are as follows:—

$\beta-$, $\beta-?-$, $\beta=$, Examiner R.

$\beta-?-$, Examiner G.

$\beta\gamma$ and $\gamma\beta$, Examiners A, E, M, P.

$\gamma\beta$, Examiner L.

We have thus a difference of several grades between the highest and the lowest limit adopted by the different examiners. In the Tables below we treat as α 's the grades from $\alpha+$ to $\alpha=$, as β 's the grades from $\beta++$ to $\beta=$; as γ 's the grades from $\gamma+$ to $\gamma-$. $\alpha\beta$ and $\beta\alpha$ are treated as borderline cases between α and β ; and $\beta\gamma$ and $\gamma\beta$ as borderline cases between β and γ .

14. We give in Tables 7 to 10 below the classification statistics of the various examiners on the foregoing basis, for the scripts marked by them.

TABLE 7
PAPER I (Ancient & Medieval History)

Mark	Examiner				
	D	K	O	P	Q
	Number of Awards				
α	2	—	3	—	—
Borderline	4	1	—	—	—
β	10	15	14	11	6
Borderline	1	—	—	2	7
γ	1	2	1	5	5
δ	—	—	—	—	—
	18	18	18	18	18
Median	β $\beta?$ + (11-12)	β + (13)	β (11)	β - (9)	$\gamma\beta$ (5)

Thus Examiner D gives two candidates clear α 's, 4 candidates a borderline mark between α and β , 10 candidates β , 1 candidate $\gamma\beta$, and 1 candidate γ . Q returns them all as β or worse, and no examiner uses δ .

15.

TABLE 8
PAPER II (Medieval and Modern History)

Mark	Examiner									
	A	B	C	F	H	J	K	L	N	R
	Number of Awards									
α	1	4	2	—	—	—	—	—	—	1
Borderline	1	2	3	4	1	—	2	—	2	3
β	14	10	12	12	9	13	12	15	10	12
Borderline	—	—	—	1	4	1	1	2	4	1
γ	—	1	—	—	3	3	2	—	—	—
δ	1	—	—	—	—	—	—	—	1	—
	17	17	17	17	17	17	17	17	17	17
Median	β + (13)	β +? (14)	β ?+ (12)	β ?+ (12)	β ?- (10)	β - (9)	β + (13)	β ?- (10)	β (11)	β + (13)

J and L mark the scripts as β or worse, C as β or better. A and N are the only ones to use δ .

16.

TABLE 9
PAPER III (Essay)

Mark	Examiner										
	A	B	F	H	J	K	L	N	Q	R	
	<i>Number of Awards</i>										
α	1	3	3	1	1	—	—	—	—	3	
Borderline β	2	1	3	1	1	1	2	—	3	3	
Borderline γ	14	9	12	13	12	14	13	13	4	11	
δ	—	—	—	2	—	1	3	3	5	—	
γ	—	5	—	1	4	2	—	1	6	1	
δ	1	—	—	—	—	—	—	1	—	—	
	18	18	18	18	18	18	18	18	18	18	
Median	$\beta\gamma+$ (12)	$\beta+$ $\beta\gamma+$ (12-13)	$\beta\gamma+$ (12)	β (11)	$\beta\gamma+$ (12)	β $\beta\gamma+$ (11-12)	β $\beta\gamma+$ (11-12)	$\beta\gamma-$ (10)	$\beta\gamma$ (6)	$\beta+$ (13)	*

N marks all the candidates as β or worse, and F returns them as β or better. A and N again are the only examiners to use δ .

17.

TABLE 10
PAPER IV (Political Theory)

Mark	Examiner									
	A	B	E	F	G	H	J	L	M	Q
	<i>Number of Awards</i>									
α	1	2	1	1	1	1	1	—	—	—
Borderline β	3	—	4	3	2	—	—	—	2	1
Borderline γ	9	12	10	12	13	10	10	12	13	6
δ	2	—	—	—	—	3	—	4	1	5
γ	—	2	1	—	—	2	5	—	—	4
δ	1	—	—	—	—	—	—	—	—	—
	16	16	16	16	16	16	16	16	16	16
Median	$\beta\gamma+$ (12)	$\beta\gamma+$ (12)	$\beta\gamma+$ (12)	β $\beta\gamma+$ (11-12)	$\beta\gamma+$ (12)	β $\beta\gamma+$ (11-12)	$\beta+$ (13)	$\beta-$ (9)	$\beta+$ (13)	$\beta\gamma$ (6)

L marks the scripts as β or worse, while F and G mark them as β or better. A is the only examiner to use δ .

18. We have the best basis for judging the differences between individual examiners if we consider the results of those who have marked three papers, *i.e.* A, B, F, H, J, K, L and Q; Examiners A, B, F, H, J find clear α quality in some papers, whereas K, L and Q never discover this quality.

Again, B, H, J, K and Q discover clear γ quality in some papers, but A, F and L do not, though A discovers δ quality in three papers. (A and N are the only examiners who award a δ .)

19. The averages (medians) of Q ($\gamma\beta$ for Paper I and $\beta\gamma$ for Paper III and Paper IV) differ fundamentally from the rest, all of which are in the range of β 's. Of these examiners, B and L may be regarded as the extremes; their averages (medians) are set out below:—

PAPER	II	III	IV
B	$\beta+?+$ (14)	$\beta+$ $\beta?+$ (13) (12)	$\beta?+$ (12)
L	$\beta?-$ (10)	β $\beta?+$ (11) (12)	$\beta-$ (9)

Q differs definitely from all the other examiners; and we get a fairer picture of the differences likely to occur in standard if we show the range of averages (medians) of the other examiners for the four papers set out below.

	PAPER			
	I	II	III	IV
Highest	(K) $\beta+$ (13)	(B) $\beta+?+$ (14)	(R) $\beta+$ (13)	(J & M) $\beta+$ (13)
Lowest	(P) $\beta-$ (9)	(J) $\beta-$ (9)	(N) $\beta?-$ (10)	(L) $\beta-$ (9)
Difference (Number of grades)	4	5	3	4

20. There is thus between these averages (medians) about four grades difference, from $\beta+$ to $\beta-$, corresponding to the familiar difference between II (i) and II (ii) of the Honours lists of some universities. We may say that there is between the standards of these examiners about half a class difference, even leaving Q out of account.

21. It is not surprising, if there are such differences between the averages (medians), that we should find much greater differences in the marking of individual scripts.

For Paper I, Table 3 shows that Candidate No. 13 was awarded α by Examiner O and $\gamma\beta$ by Examiner P, a range of 17 grades out of a possible range of 23. Q marks him $\beta\gamma$, but both D and K mark him $\alpha\beta$.

For Paper II, Table 4 shows that Candidate No. 8 gets $\alpha-$ from B and $\gamma+$ from J, a range of 16 grades, while Candidate No. 14 gets $\alpha-$ from B and $\gamma\beta$ from H, a range of 15 grades.

For Paper III, Table 5 shows that Candidate No. 9 gets α from A, and $\gamma+$ from B, a range of 18 grades; while Candidate No. 3 gets α from R and $\beta\gamma$ from Q and N, a range of 16 grades.

For Paper IV, Table 6 shows that Candidate No. 8 gets $\alpha-$ from B and $\gamma+$ from J, a range of 16 grades.

These ranges are not affected by Q's low marking. Moreover, the average ranges (again leaving Q out of account) are as follows :—

For Paper I	-	-	-	-	8 grades
For Paper II	-	-	-	-	11 grades
For Paper III	-	-	-	-	10 grades
For Paper IV	-	-	-	-	9 grades

Thus on the average there is a whole class difference or thereabouts between the marks awarded by different examiners to the same script, since each class may be supposed to comprise about eight grades.

In no case does the same script get the same mark from all the examiners. The closest approach to equality is in judging

the obviously very poor performance of Candidate No. 11 in Paper I; he gets γ from two examiners and $\gamma -$ from the other three.

22. The discrepancies between the marks awarded by the examiners which have been the subject of discussion in the preceding paragraphs may be considered to be due to two causes (1) constant differences of standard of marking on the part of examiners (2) the presence of an element of randomness in an examiner's marking.

These points are discussed in Part II above (see p. 42 *et seq.*).

APPENDIX II.

BRIEF SUMMARY OF THE WORK OF THE FRENCH INTERNATIONAL INSTITUTE EXAMINATIONS ENQUIRY (*Commission Française pour l'Enquête Carnegie, sur les examens et concours en France*).

1. The French Committee¹, who have received every assistance from the French Ministry of Public Instruction, have published a general report on French examinations, their character, the spirit by which they are inspired, and their relationship to the national system of education in the form of an *Atlas de l'enseignement en France* (in-quarto-raisin, pp. xiii, 183, 13 planches hors texte, à Paris, à la Maison du Livre, 4 Rue Félibien, 75 francs).

2. They also issued a questionnaire to some 4,000 persons with regard to certain examinations, and will publish a summary of the replies.

3. They have carried out a series of investigations on the *baccalauréat* examination, in many ways similar to the investigations described in the present pamphlet, and the results have been recorded in a volume entitled *La correction des épreuves écrites dans les examens, enquête expérimentale sur le baccalauréat* (in-quarto-raisin, à Paris, à la Maison du Livre, 4 Rue Félibien).

4. The first examination investigated by the Committee was the *baccalauréat*, because in their view this examination is both the most typical and the most important of all the French examinations. In the University of Paris alone there are about 15,000 candidates annually for the two parts of the *baccalauréat*. The examination serves both as a school-leaving examination for the lycées (both for boys and girls) and as an entrance examination to universities and to the liberal professions. "It is," says the French Committee, "an instrument of selection of what maybe called the directing classes" (*l'instrument de sélection des classes dites dirigeantes*)².

¹The personnel of the French Committee is given on page 7 above.

²It is clear from the context that the phrase "directing classes" is used here to designate not classes privileged by birth but those who actually exercise a directing influence in the social system. The phrase was used in the same sense in the Report of the Auxiliary Committee on Education of the Indian Statutory Commission (1929).

5. The two parts of the French *baccalauréat* correspond, roughly speaking, to the examinations for the School Certificate and for the Higher School Certificate in England. The first part is normally taken at the age of about 16 by pupils of the *classe de première* (formerly called the *classe de rhétorique*). The second part is normally taken a year later by pupils in two parallel classes, the *classe de philosophie* and the *classe de mathématiques*. In these classes philosophy is treated as the most important subject on the literary side, mathematics as the most important on the scientific side; but mathematics and other science subjects are taught in the *classe de philosophie*, while philosophy and other literary subjects are taught in the *classe de mathématiques*.

6. Both parts of the *baccalauréat* include a written examination and a *viva voce* examination in a number of subjects. Only those who pass on the written examination are admitted to the *viva voce*. A total aggregate of 50 per cent on the subjects of the written examination is required for a candidate to be admissible to the *viva voce* examination—it would appear, without a minimum requirement in any one subject.

7. The following summary is translated from the proofs of Chapter VIII of the volume :—

(1) Two investigations have been undertaken by the French Committee (Commission Française Carnegie) on the marking of scripts at the *baccalauréat* examination. The chief investigation was undertaken with reference to the examinations in :

Translation from Latin (<i>Version latine</i>)	} Part I of the <i>baccalauréat</i>
French Essay (<i>Composition française</i>)	
English	
Mathematics	
Philosophy	} Part II of the <i>baccalauréat</i> for pupils of the <i>classe de philosophie</i>
Physics	} Part II of the <i>baccalauréat</i> for pupils of the <i>classe de mathématiques</i>

100 scripts corresponding to each of these examinations, which had been actually written at the examinations held in July, 1930, were corrected and marked by 5 examiners

(*correcteurs*) chosen from the panel of examiners for the *baccalauréat* (the actual mark of the examiner at the *baccalauréat* examination furnishing a sixth mark).

The scripts chosen formed a sufficiently typical sample of the *baccalauréat* scripts as a whole.

A supplementary investigation was made on three French essays (*copies de composition française*), selected from those used for the principal investigation, which were corrected and marked by 76 different examiners.

(2) The maximum ranges¹ of the marks attributed to one and the same script in the first investigation by the different examiners were as follows:—

12 marks out of 20 for Latin translation	[60 per cent.]
13 marks out of 20 for French Essay	[65 per cent.]
9 marks out of 20 for English	[45 per cent.]
9 marks out of 20 for Mathematics	[45 per cent.]
12 marks out of 20 for Philosophy	[60 per cent.]
8 marks out of 20 for Physics	[40 per cent.]

The mean differences between the marks of two examiners varied from 1.88 out of 20 in Physics [*i.e.*, 9.40 per cent.] to 3.36 out of 20 in Philosophy [*i.e.*, 16.80 per cent.].² The number of the differences between two examiners equal to or higher than 5 marks out of 20 (25 per cent.) was 2.5 per cent. in Physics and 23 per cent. in Philosophy.

(3) The number of scripts which were recorded as deserving an average mark or a mark higher than the average in the opinion of some of the examiners (but not of all) was as follows:—

Latin translation	... 50 per cent. of the total number of the scripts
French Essay	70 per cent. of the total number of the scripts
English	... 47 per cent. of the total number of the scripts
Mathematics	36 per cent. of the total number of the scripts
Philosophy	... 81 per cent. of the total number of the scripts
Physics	... 50 per cent. of the total number of the scripts

¹The term 'range' is used, as in the text of this pamphlet, to denote the difference between the highest and lowest marks allotted by different examiners to the same script.

²The differences between each pair of examiners for each candidate were calculated, there being, with six examiners, 15 differences in respect of each candidate, and 1,500 for each subject.

8. For the second investigation on the French Essay three scripts, Nos. 23, 25 and 34, were selected, each of which at the original *baccalauréat* examination had been awarded 36 marks out of 80 (or 45 per cent.) and had been ranked as 24th out of a batch of 50. These three scripts were marked independently by 76 examiners. The marks for script No. 23 varied from 4 to 52, for script No. 25 from 12 to 64, and for script No. 34 from 16 to 56 out of a maximum of 80. The mean marks for the three scripts were as follows: Script No. 25—25·9; Script No. 25—40·0; Script No. 34—34·4.

9. The book contains an elaborate statistical analysis of the relations between the marks of the different examiners, from which the following may be quoted:—

After reduction by means of appropriate corrections of the scales of the different examiners to the same level of severity (by reducing to the same average) and to the same distribution (by altering the marks so that they have the same standard deviation), there still remain important differences between the results of the pairs of examiners. The correlation between the marks of two examiners was never perfect, with a value of $r = 1$, and was as low as $r = 0\cdot112$ (correlation between the marks of Examiner C and Examiner D in Philosophy for 50 scripts of women candidates). The mean correlation coefficient of all the examiners taken in pairs varies from $r = 0\cdot429$ in Philosophy (scripts of women candidates) to $r = 0\cdot888$ in Mathematics (scripts of male candidates).