STUDIES IN EDUCATIONAL TESTS. No. 3. ANALYSIS OF MARKS IN THE SCHOOL LEAVING CERTIFICATE EXAMINATION IN THE UNITED PROVINCES, INDIA, 1919.

By P. C. MAHALANOBIS AND KHAGENDRA NATH CHAKRAVARTI.

Statistical Laboratory, Presidency College, Calcutta.

Introduction

It is known that proficiency in one subject is sometimes associated with that in another, and it is desirable to take full advantage of such association in the organization of teaching. The correlation between marks in different subjects is the most objective method of studying such associations. Correlational analysis of examination marks however does not appear to have attracted in India the attention it deserves. In the present note an attempt has been made to study by such methods the effect of a common medium of examination, or of similarity in the nature of the subjects.

The statistical analysis is based on the marks in different subjects in the School Leaving Certificate Examination of the United Provinces for 1919. This examination corresponded at that time to the Matriculation Examination of some of the Indian Universities, and included four subjects, with one paper each in English, Mathematics, Vernacular and History-Geography. The number of candidates in 1919 for whom marks were placed at our disposal was 2,357, and their ages generally varied between 16 and 23 years. They belonged to nine different caste-groups altogether.

Our thanks are due to Rai Sahib Debnarayan Mukherjee, Vice-Principal, Teachers Training College, Agra, who very kindly supplied us with a copy of the Marks Register of the Examination which contained entries of marks in each subject, age, and caste of the candidates. We have restricted ourselves to the statistical analysis of the data, and we hope that Mr. Mukherjee himself will discuss the educational significance of the results in a subsequent paper.

In the first part of the paper we have studied the frequency distributions of marks in all the four subjects, using the Pearsonian family of curves for graduation. A detailed analysis of linear correlation is given in the second part, using Pearson's product moment formula. The various coefficients of partial correlation have also been fully worked out. An attempt has been made throughout to present the conclusions in a non-technical form, so that educational workers may have no difficulty in following the discussion.

How far marks in particular subjects depend upon age or caste of the candidates is a question of considerable interest and importance, and we hope to study it in a subsequent paper.

Frequency Distributions

Marks 0 to 100 were classified into 20 groups using a grouping unit of 5 marks[1]. The mean score, the standard deviation, and β_1 and β_2 were obtained in the usual way. The frequency constants are given in Table 1.

Table 1.—Frequency Constants for Marks. (U. P. Examination, 1919.)

Grouping Unit = 5 Marks. N = 2357

	English	Mathematics	Vernacular	History-Geography
Mean	39·48 ± ·15	47·62 ± ·23	54·41 ± ·14	39·94 ± ·16
Standard Deviation	10·78 ± ·11	16·90 ± ·17	10·16 ± ·10	11·19 ± ·11
Coeff. of Variability	27·31 ± ·29	35·49 ± ·39	18·67 ± ·19	28·02 ± ·30
β_1	·00001 ± 0000	·0005 ± ·0000	·0528 ± ·0281	·0022 ± ·0006
β_2	3·073 ± ·0768	2·724 ± ·0445	3·794 ± ·2599	3·1327 ± ·0845
Type	Normal	Normal	Type VII	Normal
μ_2	4·6476	11·4650	4·0010	5·0066
μ_3	+ 0·0363	−1·0685	−1·8383	−0·5232
μ_4	66·3852	358·2932	60·7386	78·4757

It will be noticed from the above table that in the case of English, Mathematics and and History-Geography, β_1 is practically zero and β_2 does not significantly differ from 3. We anticipate that the normal curve would give a good fit for these distributions. The observed and the expected values of the frequency distributions as graduated by the normal curve are given in Table 2 (English), Table 3 (Mathematics), Table 4 (History-Geography). The corresponding graphs are shown in Charts 1, 2 and 4.

[1] The ranges were taken 0—5, 6—10, and so on. The arrays were thus centred at 2·5, 8, 13, etc. Strictly speaking a correction for the first group is necessary. In view of the fact however that there were hardly any entries with zero marks we have not used any correction. It will be noticed that border-line cases were avoided in the above arrangement of the frequency classes. A better procedure would be to divide the border-line marks equally on either side.

TABLE 2.—GRADUATED AND OBSERVED DISTRIBUTION OF MARKS IN ENGLISH.

Mean = 39·4785 s = S.D = 10·7780 N = 2857

Range	Graduated by Normal Courve (f)	Observed (f')	$\frac{(f-f')^2}{f}$
(1)	(2)	(3)	(4)
0 – 5	1·63 } (7·34)	2 } (10)	0·9640
6 – 10	5·71 }	8 }	
11 – 15	19·97	28	0·4596
16 – 20	56·12	61	0·4248
21 – 25	127·84	131	0·0781
26 – 30	235·72	246	0·0483
31 – 35	351·85	364	0·4196
36 – 40	425·14	441	0·5917
41 – 45	415·87	397	0·8562
46 – 50	329·32	333	0·0411
51 – 55	211·12	186	2·9889
56 – 60	109·55	107	0·0594
61 – 65	46·03	38	1·4009
66 – 70	15·65	14	0·1740
71 – 75	4·31 }	4 }	
76 – 80	0·96 } (5·48)	2 } (6)	0·1011
81 – 85	0·21 }	0 }	
Total ...	2857	2857	χ^2 = 8·6072

c = number of cells = 14. n' = 14, P = 0·80
n' = 12, P = 0·66

TABLE. 3—GRADUATED AND OBSERVED FREQUENCY OF MARKS IN MATHEMATICS.

Mean = 47·5948 s = 16·93 N = 2857.

Range	Graduated by Normal curve (f)	Observed (f')	$\frac{(f-f')^2}{f}$
(1)	(2)	(3)	(4)
0 – 5	13·98	12	·2804
6 – 10	17·10	21	·8894
11 – 15	32·78	41	2·0613
16 – 20	57·66	67	1·5129
21 – 25	92·97	103	1·0821
26 – 30	137·48	118	2·7062
31 – 35	186·53	205	1·8289
36 – 40	231·91	231	·0188
41 – 45	264·50	264	·0008
46 – 50	276·75	262	·7861
51 – 55	265·37	263	·0211
56 – 60	233·43	232	·0088
61 – 65	188·37	190	·0141
66 – 70	139·30	132	·3825
71 – 75	94·51	85	·9569
76 – 80	58·81	68	1·4361
81 – 85	33·54	37	·3569
86 – 90	17·55	16	·1368
91 – 95	8·42 } 14·46	7	3·3833
96 – 100	6·04 }		
Total	2857	2857	χ^2 = 18·3184

c = 19, n' = 19, P = 0·44
n' = 17, P = 0·30

TABLE 4—GRADUATED AND OBSERVED FREQUENCY OF MARKS IN HISTORY AND GEOGRAPHY.

Mean = 39·9481 s = 11·185 N = 2357

Range	Graduated by Normal Curve (f)	Observed (f')	$\frac{(f-f')^2}{f}$
(1)	(2)	(3)	(4)
0 – 5	2·10 } (8·75)	5 } (13)	2·0640
6 – 10	6·65 }	8 }	
11 – 15	21·59	28	1·9081
16 – 20	57·56	56	·0422
21 – 25	126·06	146	3·1540
26 – 30	226·80	216	·5143
31 – 35	335·35	315	1·2348
36 – 40	397·59	446	5·8943
41 – 45	415·86	436	·9754
46 – 50	333·04	304	2·5322
51 – 55	224·34	206	1·4993
56 – 60	124·10	109	1·8373
61 – 65	56·41	53	·2061
66 – 70	21·06	22	·0419
71 – 75	6·46 } (8·49)	6 } (7)	·2583
76 – 80	2·03 }	1 }	
Total ...	2357	2357	χ^2 = 22·1572

c = 14, n' = 14, P = ·055

n' = 12, P = ·024

TABLE 5—GRADUATED AND OBSERVED FREQUENCY OF MARKS IN VERNACULAR.

Mean = 54·4086 s = 10·1625 N = 2357

Range	Graduated by type VII Curve (f)	Observed (f')	$\frac{(f-f')^2}{f}$
(1)	(2)	(3)	(4)
0 - 5	} 11	0 } (17)	3·272
6 - 10		1	
11 - 15		1	
16 - 20		5	
21 - 25		10 }	
26 - 30	15	20	1·667
31 - 35	40·5	43	·154
36 - 40	108	105	·083
41 - 45	214	207	·229
46 - 50	396	387	·205
51 - 55	498	490	·128
56 - 60	463	468	·054
61 - 65	319	332	·529
66 - 70	169·5	174	·118
71 - 75	75·5	77	·028
76 - 80	34	25	2·382
81 - 85	10	7	·900
86 - 90	3·5	5	·643
Total ...	2357	2357	χ^2 = 10·392

c = 14, n' = 14, P = 0·66

n' = 10, P = 0·82

In the case of Vernacular alone, the frequency curve is non-normal and is of Pearsonian Type VII. The equation to the curve is given by

$$y = y_0\left(1+\frac{x^2}{a^2}\right)^{-m} \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad (1\cdot0)$$

where

$$m = (5\beta_2-9)/2(\beta_2-3) = 6\cdot 27\ 7386 \quad \ldots \quad (1\cdot1)$$

$$a^2 = 2\mu_2\cdot\beta_2/(\beta_2-3) = 38\cdot 22\ 9187 \quad \ldots \quad (1\cdot2)$$

and

$$y_0 = N.\ \Gamma(m)/a.\ \sqrt{\Pi}.\Gamma(m-\tfrac{1}{2}) = 505\cdot 902 \quad \ldots \quad \ldots \quad (1\cdot3)$$

The theoretical curve was drawn on a large scale on graph paper, and the expected frequencies were obtained by measuring the appropriate areas on the graph.

The values of χ^2 were calculated in the usual way by squaring the difference $(f-f')$ between observed (f') and expected frequencies (f) and dividing by the expected frequency (f) and adding, that is, by adding terms of the type $(f-f')^2/f$ for each cell. The value of P, the probability of occurrence of a system of discrepancies equal to or greater than that observed, can then be obtained from tables for given values of χ^2 and of the number of cells.

In using Elderton's Table XII (Biometric Tables Part I, pp. 26-27) we may put n' = the total number of cells. A more stringent test is obtained if we take into consideration the appropriate degrees of freedom. For example, in testing the graduation by a normal curve, we have estimated both the mean and the standard deviation from the sample itself. Hence there will be a loss of two degrees of freedom, and in Elderton's Table XII we should use $n'=c-2$, where c = the number of cells. For graduation by a Type VII curve we also use the third and fourth moments, so that the loss of degrees of freedom is now four, and $n'=c-4$. Instead of Elderton's Table we may use R. A. Fisher's Table III (Statistical Methods for Research Workers, 4th edition, 1932, pp. 104-105) in which the values of χ^2 for various levels of significance are given for various values of n. It may be noted that Elderton's $n'=n+1$ corresponding to Fisher's n. In the case of a normal curve Fisher's $n= c-3$, and for a Type VII curve $n= c-5$. As already noted the interpretations are different in the two cases.

The values of χ^2 and P are gathered together in Table 6. It will be noticed that in the present case it is immaterial which particular interpretation we chose, for the fit is quite satisfactory in both cases.

TABLE 6.—GOODNESS OF FIT.

Subject	Curve	c = no. of cells	n'	P	n'	P
English	Normal	14	14	0·80	12	0·66
Mathematics	,,	19	19	0·44	17	0·80
History-Geography ...	,,	14.	14	0·055	12	0·024
Vernacular	Type VII	14	14	0·66	10	0·82

The frequency distributions of marks may be considered to be practically normal in the case of English, Mathematics and History-Geography. Scores were symmetrically distributed on either side of the average value. In other words, each group of students who were below average by a certain amount are matched by another group who were above average by the same amount.

In the case of Vernacular however the value of $\beta_1 = \cdot 0528 \pm \cdot 0281$, and $\beta_2 = 3{\cdot}794 \pm \cdot 269$ is considerably greater than 3, showing that marks are more widely scattered than in the other cases. This may very likely be ascribed to the fact that Vernacular included two different languages 'Hindi' and 'Urdu', marks for both being lumped together under one head. The average scores in the two languages were probably different, and the marks were thus distributed round two different mean values. The large value of β_2 may thus be considered to have arisen from the superposition of two distinct distributions, and the non-normality of the curve is an indication of heterogeneity of the material. The distribution is also to some extent asymmetric in the case of Vernacular ; the asymmetry however is not definitely significant. The practically symmetric character of all the distributions is, we think, reassuring from an educational point of view, for there was no special discrimination against either the above-average or the below-average candidates.

The mean score in Vernacular (54·41 ± ·14) is the highest, then comes Mathematics (47·62 ± ·23), while the means in English (39·48 ± ·15) and History-Geography (39·94 ± ·16) are practically equal.

The standard deviation (which measures the scatter of marks) was 16·19 ± ·17 in Mathematics and was distinctly higher than in other subjects for which it varied between 10·16 ± ·10 and 11·19 ± ·11. The larger value of the standard deviation in Mathematics shows that the best boys and the worst boys were more widely separated from the average type in this subject. Mathematics therefore was a better discriminant of ability then other subjects.

The coefficient of variability (100 × S.D./Mean) is lowest in the case of Vernacular, showing that the marks clustered round the average more closely in Vernacular than in other subjects.

Equivalent Standards of Examination

It is usually considered desirable to have equal or equivalent standards of examination in different subjects. The concept of 'equivalent standard' does not however appear to have been clearly defined by educationists. In the present examination the distribution of marks is practically normal in all subjects. This makes it possible to set up a simple statistical definition of equivalent standards.

The variation in the distribution of marks in different subjects may be considered to have arisen from four sets of factors : (*a*) the ability or quality of the candidates, (*b*) the efficiency of teaching, (*c*) the standard of the examination, and (*d*) random fluctuations. Since the sample is fairly large and is drawn from all over a province it will not be unreasonable to assume that (*a*) the average ability and the distribution of ability of the candidates is roughly equal in different subjects, and (*b*) the average efficiency of teaching is about the same in different subjects. All the subjects will also be equally affected by (*d*) random fluctuations. Thus as a first approximation, (*c*) the standard of the examination is left as the most important single cause of variation in the distribution of marks.*

*The efficiency of teaching in different parts of the province may of course be compared with the help of the average marks and the distribution of marks in these regions. Similarly the ability of candidates in different districts or from different types of schools may be compared. But the ability of the candidates as a whole, or the teaching efficiency of the province as a whole cannot be compared unless suitable data from other provinces are available.

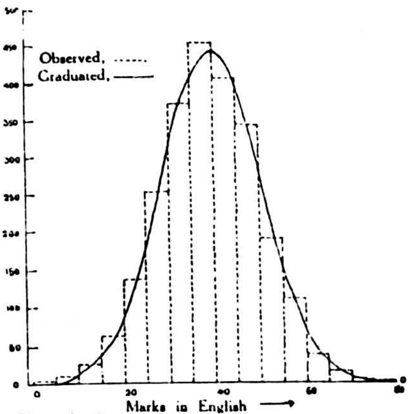

Chart 1. Frequency Distribution of Marks in English

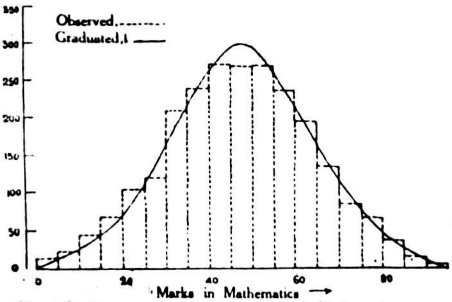

Chart 2. Frequency Distribution of Marks in Mathematics

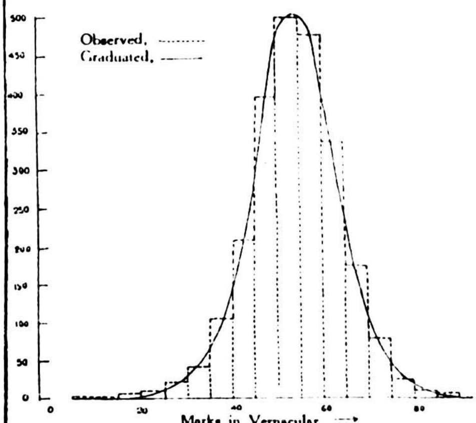

Chart 3. Frequency Distribution of Marks in Vernacular

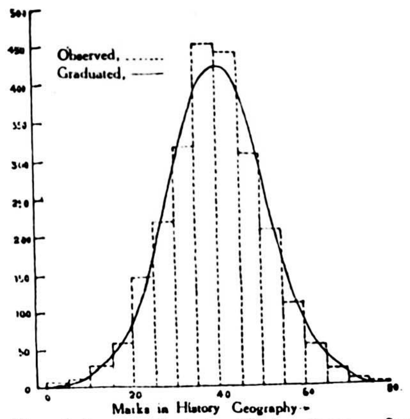

Chart 4. Frequency Distribution of Marks in History → Geography

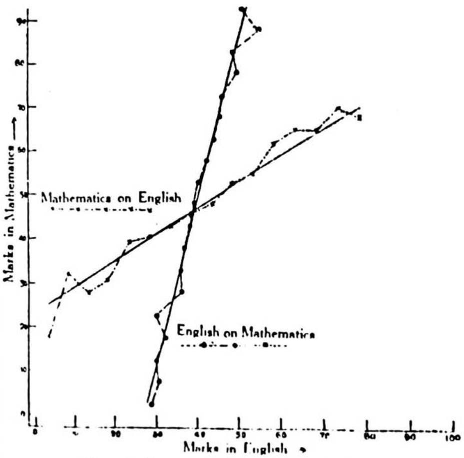

Chart 5. Regression Line– English & Mathematics

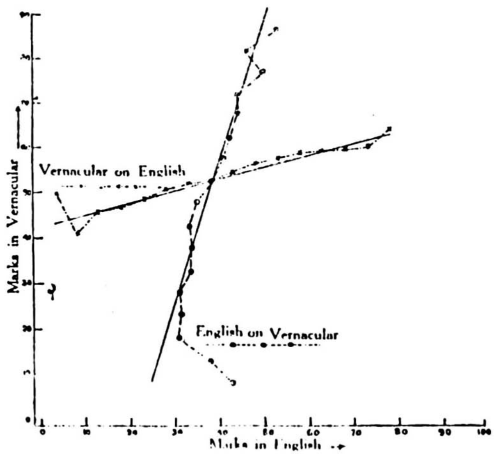

Chart 6. Regression Line– English & Vernacular

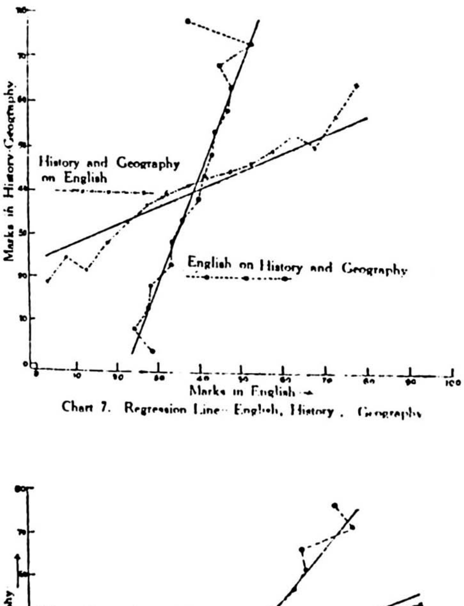

Chart 7. Regression Line– English, History , Geography

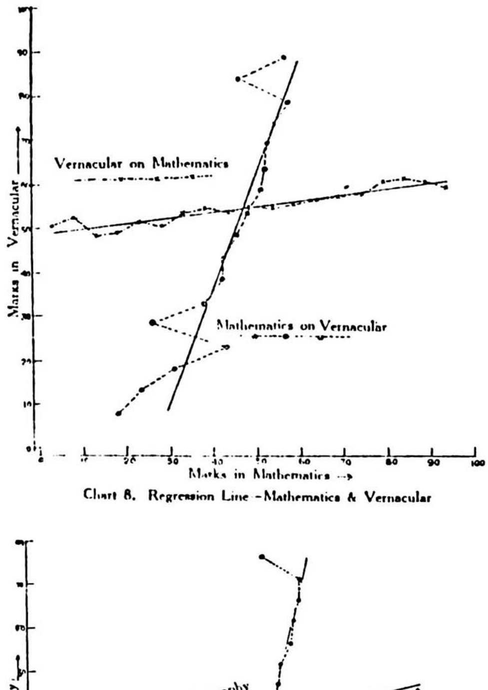

Chart 8. Regression Line –Mathematics & Vernacular

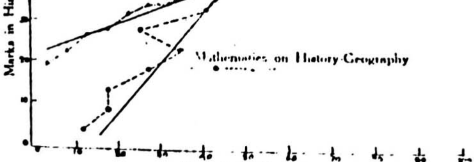

Chart 9. Regression Line Mathematics, History Geography

Chart 10. Regression Line Vernacular, History Geography

It is clear that we can define the standard of the examination in two subjects to be equal when the distribution of marks in these subjects are statistically identical. For normal distributions, it is necessary and sufficient that the two average marks and the two standard deviations should be approximately equal.

The mean marks and standard deviations, are not equal in the present case, but we may adjust the marks in one subject to make them equivalent to the marks in the other subjects.

For example, let p_1, p_2, p_3, and p_4 be the pass marks in English, Mathematics, Vernacular, and History-Geography respectively. Also let m_1, m_2, m_3, and m_4 be the corresponding mean scores, and s_1, s_2, s_3, and s_4 the corresponding standard deviations. As the distributions are practically normal, the number of failures will be given by the probability integral from 0 to $(m - p)/s$ in each case.

We may thus define equivalent standards of pass in each subject by making

$$(m_1 - p_1)/s_1 = (m_2 - p_2)/s_2 = (m_3 - p_3)/s_3 = (m_4 - p_4)/s_4$$

In the present case we have the following mean values and deviations.

TABLE 7.—MEAN VALUES AND, STANDAND DEVIATIONS.

	Mean Marks	Standard Deviation
English	$m_1 = 39·48$	$s_1 = 10·78$
Mathematics ..	$m_2 = 47·62$	$s_2 = 16·90$
Vernacular ...	$m_3 = 54·41$	$s_3 = 10·16$
History-Geography ..	$m_4 = 39·94$	$s_1 = 11·19$

If we take English as the standard, we can work out equivalent pass marks with the help of the above equation. For example, putting $p_1 = 30$, *i.e.* if we take 30 as pass marks in English we get the following values.

TABLE 8.—EQUIVALENT PASS MARKS IN DIFFERENT SUBJECTS

English (Standard) ..	30	33	36	40
Mathematics	33 (32·76)	37 (37·17)	42 (42·16)	48 (48·14)
Vernacular	45 (45·48)	48 (48·13)	51 (51·13)	55 (54·90)
History-Geography ...	30 (30·10)	33 (33·02)	36 (35·88)	40 (40·4)

We may use the same method also for fixing the limits for first, second, or third division marks.

Instead of adjusting the limits p_1, p_2, p_3 and p_4, we may keep these constants, and secure the same results by adjusting the scores for each candidate. For example, using the numerical examples given above we can construct tables of equivalent marks for Mathematics in the following way.

Let x_1 be the marks in English (standard), and let x_2 be the equivalent marks in Mathematics. Then

$$(x_1-39{\cdot}48)/10{\cdot}78 = (x_2-47{\cdot}62)/16{\cdot}90$$
$$\text{or } x_2 = 1{\cdot}568\,(x_1)-14{\cdot}3.$$

The following table shows the actual values at intervals of 10.

TABLE 9.—EQUIVALENT MARKS IN ENGLISH AND MATHEMATICS

English (x_1)	Mathematics (x_2)		Correction
	Actual	Approximate	
10	1·6	2	+ 8
20	17·0	17	+ 3
30	32·7	33	− 3
40	48·4	48	− 8
50	64·3	64	−14
60	80·8	81	−21
70	95·5	96	−26

The necessary adjustment for each mark from 0 to 100 in each subject for any given set of values of p_1, p_2, p_3 and p_4 can be easily worked out. It is clear that with such adjusted marks, mean values and standard deviations in each subject will be identical, and all the adjusted marks will be exactly comparable.

CORRELATION BETWEEN MARKS IN DIFFERENT SUBJECTS

The Coefficients of Correlation (r) between marks in different subjects are shown in Table 10. The coefficients were all calculated by the Pearson product-moment formula and the size of the sample is 2357 in each case.

TABLE 10.—TOTAL COEFFICIENTS OF CORRELATION (N=2357)

	Subjects Correlated	Symbol	Coefficient of Correlation with probable error
(1)	English (x_1) and Mathematics (x_2)	r_{12}	+ 0·3996 ± ·0117
(2)	,, ,, Vernacular (x_3)	r_{13}	+ 0·3131 ± ·0125
(3)	,, ,, History-Geography (x_4)	r_{14}	+ 0·4334 ± ·0112
(4)	Mathematics (x_2) ,, Vernacular (x_3)	r_{23}	+ 0·2304 ± ·0132
(5)	,, ,, History-Geography (x_4)	r_{24}	+ 0·5561 ± ·0096
(6)	Vernacular (x_3) ,, History-Geography (x_4)	r_{34}	+ 0·2404 ± ·0131

Writing x_1, x_2, x_3 and x_4 for marks in English, Mathematics, Vernacular, and History-Geography, and using the regression formula

$$(y-\bar{y})/s_y = r_{xy}\cdot(x-\bar{x})/s_x$$

TABLE 11.—CORRELATION CHART FOR MARKS IN ENGLISH AND MATHEMATICS (U.P.S.L.C. 1919)

Marks in English (x_1)

Marks in Mathematics (x_2)

RANGE	0—5	6—10	11—15	16—20	21—25	26—30	31—35	36—40	41—45	46—50	51—55	56—60	61—65	66—70	71—75	76—80	TOTAL
0—5	1		1	2	1	3		1		3							12
6—10			1	2	2	4	5	4	2	1							21
11—15		1	2	5	9	3	11	2	3	1	3	1					41
16—20		1	3	9	5	8	11	17	9	2		1		1			67
21—25		1	4	5	10	26	14	22	14	4	2		1				103
26—30			2	5	10	20	25	19	19	11	5		1	1			118
31—35	1	2	4	6	17	29	44	36	32	24	6	2	1		1		205
36—40		1	2	8	13	34	41	51	42	23	13	5	1				234
41—45			2	9	15	27	42	62	49	33	18	5	2				264
46—50		1		4	14	24	42	54	54	32	29	6	1	1			262
51—55		1		1	12	26	46	56	36	47	27	8	2	1			263
56—60			1	4	9	12	38	32	51	44	19	17	3	1		1	232
61—65				1	6	11	23	34	29	37	21	24	2	2			190
66—70					4	9	11	24	18	31	16	11	7	1			132
71—75					5	4	6	10	21	14	10	10	2	3			85
76—80						4	3	9	8	14	10	7	10	1	1	1	68
81—85							3	7	4	10	4	6	1	1	1		37
86—90									4	2	3	2	2	2	1		16
91—95								1	2		1	2	1				7
TOTAL	2	8	22	61	132	244	365	441	397	333	187	107	37	15	4	2	2,357

TABLE 12.—CORRELATION CHART FOR MARKS IN ENGLISH AND VERNACULAR (U.P. S.L.C. 1919)

Marks in English (x_1)

Marks in Vernacular (x_3)

RANGE	0—5	6—10	11—15	16—20	21—25	26—30	31—35	36—40	41—45	46—50	51—55	56—60	61—65	66—70	71—75	76—80	TOTAL
0—5																	
6—10									1								1
11—15								1									1
16—20				1	1			3									5
21—25			1			3	3	2	1								10
26—30		2	1	2	5	1	2	2	1		2	1	1				20
31—35		1		4	6	7	9	9	3	2	1	1					43
36—40	1		3	8	13	14	12	26	11	10	4	3					105
41—45		2	4	13	17	38	43	42	28	15	3	1		1			207
46—50		1	5	9	25	48	70	81	69	39	25	12	3				387
51—55		2	5	11	25	57	83	100	86	67	30	14	9	1			490
56—60			1	8	20	41	79	74	79	81	47	26	5	5	2		468
61—65	1		1	4	15	22	36	55	72	51	39	22	7	5	1	1	332
66—70			1	1	3	9	19	27	30	41	18	15	6	2	1	1	174
71—75					2	3	9	16	11	17	9	5	4	1			77
76—80						1		2	3	6	6	5	2				25
81—85								1	2	2	2						7
86—90										2	1	2					5
TOTAL	2	8	22	61	132	244	365	441	397	333	187	107	37	15	4	2	2,357

Table 13.—Correlation Chart for Marks in English and History-Geography (U.P. S.L.C. 1919)

Marks in English (x_1)

Marks in History and Geography (x_4)

Range	0–5	6–10	11–15	16–20	21–25	26–30	31–35	36–40	41–45	46–50	51–55	56–60	61–65	66–70	71–75	76–80	Total
0–5			1	3	1												5
6–10	1	1	1			3	1	2									9
11–15			6	4	3	1	5	5	1	2							27
16–20		2	3	9	11	5	15	5	4	1	1						56
21–25		1	2	10	21	28	23	27	16	13	3	1		1			146
26–30	1	2	6	11	27	33	45	35	24	22	8	4					218
31–35		1	2	14	22	50	74	63	43	31	12	3	1	1			317
36–40		1	1	2	17	46	63	105	89	63	34	14	4	2	1		441
41–45				5	13	39	65	89	90	69	35	19	7	1			433
46–50				1	8	18	37	55	58	66	38	21	4	2			308
51–55				2	3	14	21	27	43	38	30	20	5	3	1		206
56–60					3	4	10	14	18	15	19	16	7	1	2	1	109
61–65					8	1	4	9	9	5	3	5	8	4			58
66–70						2	2	4	2	5	3	3				1	22
71–75										3	1	1	1				6
76–80								1									1
Total	2	8	22	61	132	244	365	441	397	333	187	107	37	15	4	2	2,337

Table 14.—Correlation Chart for Marks in Mathematics and Vernacular (U.P.S.L.C 1919)

Marks in Mathematics. (x_2)

Marks in Vernacular (x_3)

Range	0–5	6–10	11–15	16–20	21–25	26–30	31–35	36–40	41–45	46–50	51–55	56–60	61–65	66–70	71–75	76–80	81–85	86–90	91–95	Total
6–10				1																1
11–15					1															1
16–20				2	1		1						1							5
21–25					1		2	2	1		3	1								10
26–30		2	2	3	2	6	1	1	1	1		1								20
31–35			5	1	2	5	7	3	5	5	3	3	3				1			43
36–40	2	1	2	7	8	5	15	12	12	8	13	7	4	4	4	1				105
41–45	1	3	8	10	11	18	20	22	30	21	18	14	16	11	2	1		1		207
46–50	3	3	8	9	18	27	34	41	45	50	46	42	25	16	14	2	3		1	387
51–55	3	5	5	14	26	17	41	57	62	50	58	52	38	30	12	14	4	2		490
56–60	2	1	5	8	17	21	34	38	47	59	57	55	43	28	20	18	8	5	2	468
61–65	1	3	4	8	8	8	29	27	39	41	37	29	27	22	19	13	7	4	4	332
66–70		2	2	4	3	6	10	18	17	20	17	10	26	11	8	9	8	3		174
71–75					4	4	6	9	4	7	11	7	8	7	4	4	6	1		77
76–80		1				1	3	2	1			7	2	2	2	4				25
81–85					1		1	2				2		1						7
86–90							1					2	2							3
Total	12	21	41	67	103	118	205	234	264	262	263	232	190	132	85	68	37	16	7	2,337

TABLE 15.—CORRELATION CHART FOR MARKS IN MATHEMATICS AND HISTORY-GEOGRAPHY (U.P.S.L.C 1919)

Marks in Mathematics (x_2)

Marks in History and Geography (x_1)

RANGE	0—5	6—10	11—15	16—20	21—25	26—30	31—35	36—40	41—45	46—50	51—55	56—60	61—65	66—70	71—74	75—80	81—85	86—90	91—95	TOTAL
0—5	1		3	1																5
6—10	1	1		2	3		2													9
11—15	4	5	3	5	4	2	2	1	1											27
16—20	1	2	7	6	6	6	15	4	3	1	4	1								56
21—25		6	4	16	13	16	21	14	14	18	15	4	3	1			1			146
26—30	2	5	9	13	22	21	33	32	29	24	14	9	3	1			1			218
31—35	3	...	7	9	24	19	40	50	39	42	34	27	10	7	2	2	2			317
36—40		2	6	6	13	24	35	51	62	61	51	52	41	14	11	9	3			441
41—45			2	4	12	19	32	47	58	47	57	56	36	27	19	12	2	2	1	433
46—50				4	4	8	16	17	29	29	45	36	42	29	21	14	6	5	3	308
51—55				...	2	3	8	11	20	22	32	25	30	21	14	10	6	2		206
56—60				1			1	4	6	8	6	17	15	20	12	9	6	3	1	109
61—65				...				2	2	6	2	3	8	10	4	8	6	2		53
66—70								1	1	4	3	1	1	2	1	3	3	1	1	22
71—75												1	1			1	1	1	1	6
76—80															1		...			1
TOTAL	12	21	41	67	103	118	205	234	264	262	263	232	190	132	85	68	37	16	7	2,357

TABLE 16.—CORRELATION CHART FOR MARKS IN VERNACULAR AND HISTORY-GEOGRAPHY. (U.P.S.L.C. 1919)

Marks in Vernacular (x_3)

Marks in History and Geography (x_4)

RANGE	0—5	6—10	11—15	16—20	21—25	26—30	31—35	36—40	41—45	46—50	51—55	56—60	61—65	66—70	71—75	76—80	81—85	86—90	TOTAL
0—5							1		2		1		1						5
6—10					1			1		1	5	1							9
11—15				1		1	1	5	9	4	4		2						27
16—20			1	1		3	1	8	7	14	7	8	3	3					56
21—25				1	2	3	6	9	16	20	35	30	15	5	3	1			146
26—30						4	9	16	19	46	43	36	24	13	6	2			218
31—35				1		3	8	13	43	68	65	56	33	18	7	1	1		317
36—40				1	3	2	4	20	41	78	97	86	61	26	13	5	3	1	441
41—45					2	3	6	14	30	74	84	98	61	33	22	4		2	433
46—50		1				1	5	14	18	41	66	60	58	31	12	3	2	1	308
51—55					1	...	1	3	17	24	48	45	41	16	7	2	1		206
56—60					1		1	1	2	10	19	30	22	14	6	3			109
61—65								1	3	5	8	13	11	7		4		1	53
66—70										1	6	4	4	6	1				22
71—75											2	1	1	2					6
76—80										1									1
TOTAL		1	1	5	10	20	43	105	207	387	490	468	332	174	77	25	7	5	2,357

we get the following regression equations.

$$x_1 = 27{\cdot}37 + 0{\cdot}25\ (x_2)$$
$$x_1 = 21{\cdot}41 + 0{\cdot}33\ (x_3)$$
$$x_1 = 22{\cdot}80 + 0{\cdot}42\ (x_4)$$

$$x_2 = 22{\cdot}81 + 0{\cdot}63\ (x_1)$$
$$x_2 = 26{\cdot}71 + 0{\cdot}38\ (x_3)$$
$$x_2 = 13{\cdot}97 + 0{\cdot}84\ (x_4)$$

$$x_3 = 42{\cdot}75 + 0{\cdot}30\ (x_1)$$
$$x_3 = 47{\cdot}83 + 0{\cdot}14\ (x_2)$$
$$x_3 = 45{\cdot}68 + 0{\cdot}22\ (x_4)$$

$$x_4 = 22{\cdot}19 + 0{\cdot}45\ (x_1)$$
$$x_4 = 22{\cdot}46 + 0{\cdot}37\ (x_2)$$
$$x_4 = 25{\cdot}55 + 0{\cdot}26\ (x_3)$$

The observed mean values of arrays in Tables 11—16 are given in Table 17, and are shown in Charts 5—10 together with the corresponding regression lines.

Two-Factors Analysis.

In the Two-Factors theory of abilities developed by Prof. Spearman*, each mental ability may sometimes be resolved into two factors, one of which is usually called '*g*' and is supposed to be common to all abilities, and the other '*s*' is specific to each ability. A necessary consequence of this theory is that there must exist certain algebraic connexions between the magnitude of the different coefficients of correlation.

Let x_i, x_j, x_k, x_l .... represent marks in different subjects (or scores in performance in tests of different abilities), σ_i, σ_j, σ_k, σ_l .... the corresponding standard deviations, and ϱ_{ij}, ϱ_{ik}, ϱ_{kl} .... etc. the coefficients of correlation in the parent population. Spearman's Two-Factors theory then requires that tetrad equations of the following type $(\varrho_{ij} \cdot \varrho_{kl} - \varrho_{ik} \cdot \varrho_{jl})$ must vanish for all possible combinations of four subjects or abilities.

In actual practice we cannot work with the hypothetical parent populations, but must be satisfied with samples drawn from such parent populations. In other words instead of working with σ_i, σ_j σ_k .... and ϱ_{ab}, ϱ_{cd} .... etc. we must work with the observed parameters in samples s_i, s_j .... r_{ij}, r_{ik}, r_{kl} .... etc. Now even when the tetrads vanish in the parent population, *i.e.* $\varrho_{ij} \cdot \varrho_{kl} - \varrho_{ik} \cdot \varrho_{jl} = 0$ the corresponding sample values $r_{ij} \cdot r_{kl} - r_{ik}\ r_{jl}$ will not be all arithmetically equal to zero. Therefore in order to judge whether they may be considered statistically negligible or zero, it is necessary to use expressions for the standard (or probable) errors of tetrads.

There are, however, certain difficulties in estimating the variance of tetrads, and it is more convenient to work with corresponding product-sums.

---

*C. Spearman—The Abilities of Man (*Macmillan & Co., Ltd., London*, 1927) gives a general exposition of the theory.

TABLE 17.—MEAN SCORE IN DIFFERENT SUBJECTS CORRESPONDING TO SCORES IN

Range	(1) English			(2) Mathematics			(3) Vernacular			(4) History-Geography		
	Mathematics	Vernacular	History-Geography	English	Vernacular	History-Geography	English	Mathematics	History-Geography	English	Mathematics	Vernacular
0-5	18·00	50·50	18·00	28·45	50·05	19·70	43·00	...		28·00	12·00	47·00
6-10	32·40	41·15	24·25	31·35	52·05	22·05	38·00	18·00	48·00	24·15	18·62	48·00
11-15	28·25	46·87	20·96	30·45	48·00	25·95	31·00	23·00	18·00	27·45	18·18	43·95
16-20	31·70	47·05	27·40	32·35	48·80	27·95	31·50	31·00	25·00	28·00	28·89	47·40
21-25	39·90	49·70	32·70	30·10	51·10	31·25	31·00	42·50	36·50	33·35	35·57	51·46
26-30	40·75	51·15	36·10	35·55	50·75	33·90	33·70	26·50	30·00	33·95	25·92	52·29
31-35	43·68	52·80	37·35	35·65	53·05	34·10	34·30	38·95	33·70	35·70	41·57	52·59
36-40	46·47	53·15	39·65	37·10	54·15	37·30	33·75	41·25	34·45	40·05	47·18	54·33
41-45	48·74	55·40	41·85	38·25	53·45	39·15	35·05	42·05	36·20	40·95	50·07	55·45
46-50	53·10	57·80	42·85	39·60	54·65	40·15	38·65	45·85	38·10	43·20	55·61	56·05
51-55	54·70	58·53	45·40	40·20	54·40	41·20	41·15	46·99	39·90	44·46	56·64	56·40
56-60	61·60	59·60	48·33	41·95	55·54	43·04	42·65	50·40	41·35	47·10	63·18	58·85
61-65	64·80	60·40	52·05	43·70	56·20	45·85	44·55	51·05	42·70	48·47	66·11	59·51
66-70	64·75	60·65	49·30	44·52	56·75	48·53	44·25	52·51	44·10	45·96	64·81	60·50
71-75	70·50	61·75	56·75	44·70	57·55	48·75	50·20	53·15	42·49	53·00	77·16	60·50
76-80	68·00	65·50	63·00	49·30	60·60	50·40	46·55	56·80	46·00	38·00	78·00	52·00
81-85	...	...	...	48·10	61·51	52·05	53·00	45·15	36·00	...	..	...
86-90	...	...	...	54·85	60·50	54·55	...	55·00	47·00	...	...	...
91-95	...	...	...	50·85	59·35	55·10	...	...	...	...	...	...
96-100	...	...	...	..	...	...	...	..	..	...	..	...

Let $\alpha_{ij} = S(x - \bar{x}_i)(x - \bar{x}_j) = \sigma_i \sigma_j \; \rho_{ij}$

Then corresponding to the tetrad $T_{ijkl} = \rho_{ij}\rho_{kl} - \rho_{ik}.\rho_{jl}$

we have $\Pi_{ijkl} = \alpha_{ij} \; _{kl} - \; _{ik}.\alpha_{jl}$

In samples of size N, we shall then have

$$P_{ijkl} = \{\alpha_{ij}\alpha_{kl} - \alpha_{ik}\alpha_{jl}\}.\; N^2/(N-1)\;(N-2) = \sigma_i\sigma_j\sigma_k\sigma_l\;(\rho_{ij}\rho_{kl} - \rho_{ik}\rho_{jl}).\; N^2/(N-1)\;(N-2)$$
$$= \sigma_i.\sigma_j.\sigma_k.\sigma_l.\; T_{ijkl}.\; N^2/(N-1)\;(N-2) = O$$

The factor $N^2/(N-1)\;(N-2)$ is used here in order to correct for the bias introduced in basing the estimates on samples of finite size. In fact the mean value of P_{ijkl} is equal to the population value of P_{ijkl}. That is $\bar{P}_{ijkl} = \Pi_{ijkl}$.

Dr. John Wishart ("Sampling Errors in the Theory of Two-Factors", *Brit. Jour. Psychol.* Vol. 19, 1928-29, pp. 180—187) has given the following expression for the variance of P_{ijkl} in samples of size N.

$$(N-2).\sigma_P^2 = \sigma_1^2\sigma_2^2\sigma_3^2\sigma_4^2\;[\rho_{13}^2 + \rho_{14}^2 + \rho_{23}^2 + \rho_{24}^2] - 2\;(\rho_{12}\rho_{13}\rho_{23} + \rho_{12}\rho_{14}\rho_{24} + \rho_{13}\rho_{14}\rho_{34} + \rho_{23}\rho_{24}\rho_{34}$$
$$+ 2\;\rho_{12}\rho_{34}\;(\rho_{13}\rho_{24} + \rho_{14}\rho_{23}) + 2\;(\rho_{13}\rho_{24} - \rho_{14}\rho_{23})^2 + 2\;(1-\rho_{12}^2)\;(1-\rho_{34}^2)/(N-1)]$$

The chief difficulty in using the above equation arises from the fact that the values of σ and ρ in the parent population are not known, and we are obliged to substitute the values of s and r as observed in the sample itself. In the present state of our knowledge this procedure is however unavoidable.

The actual values of P and s_P^2 are given below in Table 8. The corresponding value of the tetrads T and s_T^2 are also given, it being assumed that s_T^2 will be given by $s_P^2/s_1^2s_2^2s_3^2s_4^2$

TABLE 18.—VALUES OF P AND T (TETRAD EQUATIONS)

Observed P		s_P^2	Observed Tetrad T		s_T^2
Symbol	Value		Symbol	Value	
$a_{13}a_{24} - a_{14}a_{23}$	2·4697	·8956	$r_{13}r_{24} - r_{14}r_{23}$	·0743	·0112
$a_{12}a_{34} - a_{13}a_{24}$	2·5927	·8828	$r_{12}r_{34} - r_{13}r_{24}$	·0780	·0115
$a_{14}a_{23} - a_{12}a_{34}$	0·1280	·2652	$r_{14}r_{23} - r_{12}r_{34}$	·0037	·0080

*Prof. Karl Pearson and Miss Margaret Moul have discussed certain aspects of this question in great detail in "Mathematics of Intelligence. I. Sampling Errors in the Theory of a Generalized "Factor." *Biometrika*, Vol. 19, 1927, pp. 246-291.

It is clear from Table 18 that we can consider only one of the equations, *e.g.* $r_{14}\, r_{23} - r_{12}\, r_{34}$ to have vanished within the limits of accuracy of sampling errors. The other two equations are significantly different from zero, and hence the tetrad conditions are not satisfied. We must conclude therefore that a simple analysis into "Two Factors' is not possible in the present case. Apart from a general factor and specific factors, a certain amount of overlapping between proficiency in different subjects is indicated by the above analysis.

Factors affecting the correlation

We may therefore not unreasonably consider the correlations in Table 10 to have arisen from one or other of the following factors or a combination of them :—

(*a*) General intelligence or ability of the candidates, an able student doing more or less well in different subjects.

(*b*) Specific ability in particular subjects.

(*c*) Similarity in the nature of the subjects, training in one subject partly enforcing training in another.

(*d*) Common medium of expression; in this case, English in all subjects other than Vernacular.

Total Correlations.

In Table 10 we find that the two lowest coefficients are those between Vernacular and Mathematics $(+0.2304 \pm .0132)$, and between Vernacular and History-Geography $(+0.2404 \pm .0131)$. This low value may probably be partly ascribed to the absence of a common medium of examination and also to the absence of similarity in the nature of the subjects. The next higher correlation between Vernacular and English is distinctly larger $(+0.3131 \pm .0125)$. Although the medium is different in this case, both are language subjects, and some overlapping of training or ability may be presumed. The considerably larger correlation between English and Mathematics $(+0.3936 \pm .0117)$ may probably be ascribed partly to the common medium of examination, especially in the case of Geometry which requires exposition in English. The next higher value between English and History-Geography $(+0.4334 \pm .0112)$ is also probably due to the same factor of a common medium. Finally, in the highest correlation between Mathematics and History-Geography $(+0.5561 \pm .0096)$ we probably find the effect of a real similarity in the subject matter in the case of Geography and Mathematics besides the effect of a common medium.

Partial Correlation.

The coefficients of partial correlations were then calculated in accordance with the following formula :—

$$r_{12\cdot3} = \frac{r_{12} - r_{13}\cdot r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}}$$

All the coefficients of partial correlation of the first order, twelve altogether, are shown in the following Table 19.

TABLE 19.—PARTIAL COEFFICIENTS OF THE FIRST ORDER.

	SUBJECTS CORRELATED	Subject eliminated	Symbol	Correlation coefficient
1	English (x_1) and Mathematics (x_2) ..	Vernacular (x_3) ..	$r_{12.3}$	+0·3533
2	English (x_1) and Mathematics (x_2) ..	History-Geography (x_4) ...	$r_{12.4}$	+0·2117
3	English (x_1) and Vernacular (x_3) ...	Mathematics (x_2) ...	$r_{13.2}$	+0·2478
4	English (x_1) and Vernacular (x_3) ..	History-Geography (x_4) ..	$r_{13.4}$	+0·2388
5	English (x_1) and History-Geography (x_4) ..	Mathematics (x_2) ...	$r_{14.2}$	+0·2772
6	English (x_1) and History-Geography (x_4) ..	Vernacular (x_3) ..	$r_{14.3}$	+0·3885
7	Mathematics (x_2) and Vernacular (x_3) ...	English (x_1) ..	$r_{23.1}$	+0·1209
8	Mathematics (x_2) and Vernacular (x_3) ...	History-Geography (x_4) ..	$r_{23.4}$	+0·1199
9	Vernacular (x_3) and History-Geography (x_4)	English (x_1) ..	$r_{34.1}$	+0·1223
10	Vernacular (x_3) and History-Geography (x_4)	Mathematics (x_2) ..	$r_{34.2}$	+0·1389
11	Mathematics (x_2) and History-Geography (x_4)	English (x_1)	$r_{24.1}$	+0·4635
12	Mathematics (x_2) and History-Geography (x_4)	Vernacular (x_3) ...	$r_{24.3}$	+0·5301

The total correlation between English and Mathematics +0·3996 is reduced to +0·3533 when the influence of Vernacular was eliminated and to +0·2117 when the influence of History-Geography was eliminated. This indicates that Vernacular is much less closely associated with English and Mathematics than History-Geography. This may be attributed to the common medium in the three subjects English, History-Geography and Mathematics (including Geometry which requires exposition in English).

The total correlation between English and Vernacular came down from +0·3131 to +0·2478 and +0·2388 respectively when the effects of Mathematics and History-Geography were eliminated. As these coefficients are not significantly different from each other, we conclude that both Mathematics and History-Geography have practically equal effect on the correlation between English and Vernacular.

The correlation between English and History-Geography is reduced from +0·4334 to +0·2772 and +0·3885 respectively when the influence of Mathematics and Vernacular were eliminated. Mathematics evidently exerts a greater influence than Vernacular on the correlation between English and History-Geography.

The correlation between Vernacular and Mathematics (+0·2304), and between Vernacular and History-Geography (+0·2404) are practically reduced equally to +0·1209 and +0·1223 respectively when English was eliminated and to +0·1199 and +0·1389 when Mathematics was eliminated.

The rather high correlation between Mathematics and History-Geography (+0·5561) is on the other hand only very slightly reduced to +0·5301 on eliminating Vernacular, and to a larger extent to +0·4635 when English is eliminated. The larger reduction in the case of English is again very likely due to the influence of a common medium.

PARTIAL CORRELATIONS OF THE SECOND ORDER

Finally the coefficients of partial correlation of the second order were calculated by the formula

$$r_{12\cdot34} = \frac{r_{12\cdot4} - r_{13\cdot4}\ r_{23\cdot4}}{\sqrt{(1 - r_{13\cdot4}{}^2)\ (1 - r_{23\cdot4}{}^2)}}$$

and the results given in Table 20 were obtained.

TABLE 20.—PARTIAL CORRELATIONS OF THE SECOND ORDER.

	SUBJECTS CORRELATED (Influence of the other two subjects being eliminated)	Symbol	Coefficient of Correlation
1	English (x_1) and Mathematics (x_2) ...	$r_{12\cdot34}$	+0·1899
2	English (x_1) and Vernacular (x_3)	$r_{13\cdot24}$	+0·2199
3	English (x_1) and History-Geography (x_4)	$r_{14\cdot23}$	+0·2586
4	Mathematics (x_2) and Vernacular (x_3)	$r_{23\cdot14}$	+0·0781
5	Mathematics (x_2) and History-Geography (x_4) ...	$r_{24\cdot13}$	+0·4557
6	Vernacular (x_3) and History-Geography (x_4)	$r_{34\cdot12}$	+0·0754

The above Table gives us the coefficients of correlations between any pair of subjects, when the influence of the remaining two subjects is eliminated.

The residual correlation between English and Mathematics (+0·1899) or between English and Vernacular (+0·2199) are of the same order, while the residual correlation between English and History-Geography (+0·2534) is considerably larger. Apart from the general intelligence or ability of the candidates, these coefficients probably represent the influence of overlapping in subject-matter or training or else the effect of a common medium.

The large reduction from +0·3996 to +0·1899 or to less than half the initial value of the correlation in the case of English and Mathematics shows that History-Geography and to a much smaller degree Vernacular exert an appreciable influence in this case.

The residual correlation between English and History-Geography (+0·2536) is quite appreciable, and is the second highest in the list. It shows that proficiency in English is to some extent associated with proficiency in History-Geography. It may be noted that in the case of English and History-Geography, it is not only that the common medium in English but both the subjects require ability in expression.

The residual correlation between Mathematics and History-Geography is only slightly affected by the elimination of English and Vernacular and is reduced from +0·5561 to +0·4557. This is the highest correlation in Table 20, and shows that proficiency in Mathematics and History-Geography generally go together.

Finally it will be noticed that the residual correlation between Vernacular and Mathematics (+0·0731) or between Vernacular and History-Geography (+0·0754) are quite small. This is probably due to the fact that neither the medium of examination is common nor is there any obvious overlapping in subject matter.

MULTIPLE CORRELATION

The multiple correlation coefficients were worked out in accordance with the formula:—

$$1 - R_{1.234}^2 = (1 - r_{12}^2)(1 - r_{13.2}^2)(1 - r_{14.23}^2)$$

and the following results were obtained

$$R_{1.234} = 0.5119$$
$$R_{2.134} = 0.5865$$
$$R_{3.124} = 0.3409$$
$$R_{4.123} = 0.6048$$

We notice that $R_{4.123}$ (·6048) and $R_{2.134}$ (·5865) are comparatively large, $R_{1.234}$ (·5119) is distinctly smaller, while $R_{3.124}$ (·3409) is very low. This shows that marks in History-Geography (x_4) and in Mathematics (x_2) are more closely connected with proficiency in other subjects than English (x_1) or Vernacular (x_3), the latter subject being the most independent of all.

The multiple regression equations can be written in the form:—

$$x = b_{12.34}\cdot(x_2) + b_{13.24}\cdot(x_3) + b_{14.23}\cdot(x_4)$$

where $b_{12.34} = r_{12.34}\cdot s_{1.34}/s_{2.34}$ etc.

and $s_{1.34}^2 = s_1^2 .\ (1 - r_{13}^2) .\ (1 - r_{14.3}^2)$

$s_{2.34}^2 = s_2^2 .\ (1 - r_{23}^2) .\ (1 - r_{24.3}^2)$ etc.

The following equations were obtained:—

$$x_1 = 11{\cdot}48 + 0{\cdot}18\,(x_2) + 0{\cdot}21\,(x_3) + 0{\cdot}26\,(x_4)$$
$$x_2 = 2{\cdot}89 + 0{\cdot}28\,(x_1) + 0{\cdot}10\,(x_3) + 0{\cdot}70\,(x_4)$$
$$x_3 = 39{\cdot}75 + 0{\cdot}23\,(x_1) + 0{\cdot}05\,(x_2) + 0{\cdot}08\,(x_4)$$
$$x_4 = 12{\cdot}37 + 0{\cdot}24\,(x_1) + 0{\cdot}30\,(x_2) + 0{\cdot}07\,(x_3)$$

It will be noticed that marks in Vernacular are not dependent to any appreciable extent upon those in Mathematics and History-Geography but they are dependent to some extent on English, a score of 10 marks in English being associated on an average with a gain of 2·3 marks in Vernacular. The independence of marks in Vernacular is also shown by the fact that the average score of candidates securing zero in the three other subjects is so high as 39·75.

Scores in Mathematics are associated appreciably with scores in English, and especially with scores in History-Geography, a score of 10 marks in the latter subject leading to an average gain of 7 marks in Mathematics. The dependence of proficiency in Mathematics on proficiency in other subjects is shown by the result that a candidate scoring zero in other subjects will secure only 2·9 marks in Mathematics.

Scores in English or History-Geography are dependent on those in the other subjects almost to the same extent ; a candidate getting zero in the remaining three subjects will secure 11·5 and 12·4 marks respectively in English and History-Geography. In the case of English, the scores depend to a greater extent upon those in History-Geography than in any other subject, while in the case of History-Geography they depend more on Mathematics than on English, while the effect of Vernacular is almost negligible.

Summary of Conclusions

In the present paper we have given a statistical analysis of marks obtained by 2357 candidates in four subjects, English, Mathematics, Vernacular, and History-Geography in the U. P. School Leaving Certificate Examination in 1919. The chief results of the analysis are given below.

(1) The distribution of marks was practically symmetrical in all four subjects showing that there was no special discrimination against either candidates who were above the average in ability or against candidates who were below average in ability.

(2) The frequency distribution of marks in English, Mathematics, and History-Geography can be adequately described by the normal curve of errors.

(3) The distribution of marks in Vernacular shows some deviation from normality suggesting the existence of heterogeneity. This may probably be explained by the fact that marks in Vernacular did not represent marks in one single language, but included marks in both Hindi and Urdu.

(4) The mean score is highest in Vernacular ($54{\cdot}41 \pm {\cdot}14$), then comes Mathematics ($47{\cdot}62 \pm {\cdot}23$), while mean marks in English ($39{\cdot}48 \pm {\cdot}15$) and History-Geography ($39{\cdot}94 \pm {\cdot}16$) are practically equal.

(5) The standard deviation is $16{\cdot}90 \pm {\cdot}17$ in Mathematics and is distinctly higher than in other subjects for which it varies between $10{\cdot}16 \pm {\cdot}10$ and $11{\cdot}19 \pm {\cdot}11$. The candidates were thus more widely separated in this subject, showing that Mathematics was a better discriminant of ability than other subjects.

(6) In view of the practically normal distribution of marks it is possible to set up equivalent standards in different subjects such that percentages of passes or percentages attaining any assigned standard (first division, second division etc.) would be approximately equal in all subjects.

(7) There exists appreciable correlations between marks in different subjects showing that proficiency in one subject is associated with proficiency in other subjects.

(8) The relation between marks in two subjects is approximately linear in character, so that a gain of a single mark in one subject is usually associated with a proportionate gain in other subjects.

(9) The coefficients of correlation cannot however be analysed in terms of Spearman's theory of Two Factors, as all the tetrads do not vanish.

(10) The failure of the Two Factors analysis suggests that the following factors were probably at work:—

(*a*) General intelligence or ability of the candidates, an able student doing more or less well in different subjects.

(*b*) Specific ability in particular subjects.

(*c*) Similarity in the nature of different subjects, training in one subject enforcing training in another.

(*d*) Common medium of expression, in this case English in all subjects other than Vernacular.

(11) A study of the partial correlations (in which the effect of association with other subjects is eliminated) generally supports the above interpretation.

(12) The residual correlation after eliminating the influence of languages is highest between Mathematics and History-Geography (+0·4557), showing possibly some kind of similarity in subject matter probably in Mathematics and Geography.

(13) The residual correlation between English and History-Geography (+0·2536) is also quite appreciable, and is probably due to the influence of a common medium of examination and also the requirement of powers of expression.

(14) The residual correlation between English and Vernacular is +0·2199, and may probably be due to both being language subjects.

(15) The residual correlation between English and Mathematics is comparatively low (+0·1899).

(16) The association between Vernacular and Mathematics (+0·731), and between Vernacular and History-Geography (+0·0754) become negligibly small when the influence of other subjects is eliminated. This is probably due to the absence of a common medium of examination as well as the absence of any obvious overlapping in subject matter.

(17) A study of the multiple correlation coefficients show that scores in Vernacular are to a great extent independent of scores in the other subjects, scores in English are moderately associated, while scores in Mathematics and in History-Geography depend to a great extent on proficiency in the other subjects.

(18) The multiple regression equations have also been given, and exhibit clearly the nature of the interdependence between marks in various subjects. In the case of Vernacular the average mark of a candidate getting zero marks in other subjects will be no less than 39·75, in English 11·48, in History-Geography 12·37 and in Mathematics only 2·89. The gains in marks also vary widely; gains in Mathematics or History-Geography have practically no effect in the case of Vernacular while a gain of 10 marks in History-Geography is associated with an increase of 7 marks in Mathematics.

(*February, 1934*).