"Mathematical Contributions to the Theory of Evolution.—On Homotyposis in Homologous but Differentiated Organs." By KARL PEARSON, F.R.S., University College, London. Received January 20,—Read February 19, 1903.

(1.) In the paper on "Homotyposis in the Vegetable Kingdom,"* I defined homotypes as "undifferentiated like organs." In the course of that paper, I endeavoured to indicate that I was not unconscious of the influence of age, local environment, and position upon organism in modifying homotypic correlation. The object of my memoir, however, was to obtain some general appreciation of the average intensity of individuality in living forms, and to see if it approached the average value of fraternal heredity in plant or animal life. For this purpose I selected such material as was readily available, indicating the series where I thought differentiation of a sensible amount was present owing to the age, the situation, or the environment factors.

From the standpoint of theory, however, we are not compelled to adopt a mere indication of this kind. As soon as we can correlate between: (*a*) age and the quantitative character of the homologous organs, (*b*) situation on the organism and this same character, or (*c*) local environment and the character, we can allow for the differentiation of homologous parts, or reduce them to pure homotypes. In other words, homotyposis can be deduced from differentiated homologous parts, if we correct for the differentiation due to (*a*), (*b*) or (*c*). The test for the existence of such differentiation is simply the presence or absence of the corresponding correlation.

We have accordingly the following problems to find solutions for:—

(i.) To find the correction to be made to the apparent homotypic correlation, when the pairs of homologous parts are differentiated from each other by their periods of growth.

(ii.) To find the correction to be made to the apparent homotypic correlation, when each pair of homotypes is differentiated by a common period of growth from other pairs of homotypes.

(iii.) To find the correction to be made to the apparent homotypic correlation when the pairs of homologous parts are differentiated from each other by situation on the organism.

(iv.) To find the correction to be made to the apparent homotypic correlation when each pair of homotypes is differentiated by the environment of its organism from other pairs of homotypes.

It will be seen that in problems (ii) and (iv) we are dealing with true homotypes, but that the homotypic factor requires modifying for the influence of age or environment on the organism. In (i) and (iii)

* 'Phil. Trans.,' A, vol. 197, pp. 285—379.

we are not dealing with homotypes at all, but with homologous parts, and we wish to reduce them to homotypes by correcting for differences between them due to growth or to situation on the organism.

I propose at present to deal only with problems (i) to (iii), not because (iv) does not admit of theoretical treatment, but because we have not thus far obtained data to illustrate satisfactorily the correlation between character and the immediate environment of the individual organism. Experimental determinations of homotyposis in plants, when the individuals are subjected to a graduated environmental scale, *e.g.*, in depth of soil or quantity of moisture allowed would be fairly easy to carry out, and most interesting in result. I hope it may be possible to arrange experiments of this kind for the coming summer. We can then illustrate the fourth proposition from actual observation, and the publication of its theoretical solution will be of greater value.

(2.) *To find the correction to be made to the apparent homotypic correlation when the pair of homologous parts are differentiated from each other by their periods of growth.*

Let x and y denote the characters in the two homologous parts quantitatively determined, and t_1, t_2 their respective periods of growth. Then we have four variable quantities x, y, t_1, t_2, no one of which fixes absolutely any other, for individuals will have different characters even with the same period of growth. The proposition accordingly reduces to this: What is the correlation R between x and y for constant values of the variables, *i.e.*, selected values of, t_1 and t_2?

This problem is answered in formulæ (lviii), (lix) and (lx) of my memoir: "On the Influence of Natural Selection on the Variability and Correlation of Organs."*

Let us write in those formulæ t_1 for the subscript 1, t_2 for 2, x for 3, and y for 4; we have at once

$$\Sigma_x^2 = \sigma_x^2 \frac{1 - r_{t_1t_2}^2 - r_{xt_1}^2 - r_{xt_2}^2 + 2r_{t_1t_2}\, r_{xt_1}\, r_{xt_2}}{1 - r_{t_1t_2}^2} \ldots\ldots\ldots\ldots \quad \text{(i)},$$

$$\Sigma_y^2 = \sigma_y^2 \frac{1 - r_{t_1t_2}^2 - r_{yt_2}^2 - r_{yt_1}^2 + 2r_{t_1t_2}\, r_{yt_2}\, r_{yt_1}}{1 - r_{t_1t_2}^2} \ldots\ldots\ldots\ldots \quad \text{(ii)},$$

$$\Sigma_x\Sigma_y\text{R} = \sigma_x\sigma_y \frac{r_{xy}(1 - r_{t_1t_2}^2) - r_{xt_1}\, r_{yt_1} - r_{xt_2}\, r_{yt_2} + r_{t_1t_2}(r_{xt_1}\, r_{yt_2} + r_{yt_1}\, r_{xt_2})}{1 - r_{t_1t_2}^2} \quad \text{(iii)}.$$

Now if we deal with direct and not cross-homotyposis, *i.e.*, with the *correlation* of the *same* character in two homologous parts, we can put these results more simply. We in this case render our correlation tables symmetrical by entering each one of a pair of homologues first as an x and then as a y. We may then write

* 'Phil. Trans.,' A, vol. 200, p. 30.

$$r_{xt_1} = r_{yt_2} = r, \qquad r_{xt_2} = r_{yt_1} = r',$$
$$r_{xy} = \rho, \qquad r_{t_1t_2} = \mathbf{r}, \qquad \sigma = \sigma_x = \sigma_y,$$

and we find

$$\Sigma_x{}^2 = \Sigma_y{}^2 = \sigma^2 \frac{1 - \mathbf{r}^2 - r^2 - r'^2 + 2\mathbf{r}rr'}{1 - \mathbf{r}^2}$$

$$\mathrm{R} = \rho \frac{1 - \mathbf{r}^2}{1 - \mathbf{r}^2 - r^2 - r'^2 + 2\mathbf{r}rr'} - \frac{2rr' - \mathbf{r}(r^2 + r'^2)}{1 - \mathbf{r}^2 - r^2 - r'^2 + 2\mathbf{r}rr'} \ldots\ldots \text{(iv).}$$

This is the full solution of the first problem.

We see that in order to solve it, it is necessary:

(i.) To find the correlation ρ of the homologous pairs as if they were simple homotypes.

(ii.) To find the correlation $\mathbf{r}$ between the growth periods of each pair of homotypes.

(iii.) To find the correlation r between the character and the period of growth.

(iv.) To find the correlation r' between the character of one homotype and the period of growth of its fellow.

Now these correlations can be found at once by the usual statistical processes, if the data are forthcoming.

(3.) I propose to illustrate this on material, which, although not homotypic, is so analogous that it brings out all the important features.

We will determine the correlation between the head-length of brothers, such length being measured on school boys of all ages, from 4 to 19.* It will be clear that we have here all the difficulties of the homotypic problem—resemblance due to common origin obscured by differences in the period of growth of each individual.

Table I gives the correlation of pairs of brothers without regard to their differences of age.

Table II gives the correlation between age and length of head in the same individual.

Tables IIIA and IIIB gives the correlation between the age of one brother, and the length of head of the second.

Table IV gives the correlation between the ages of pairs of brothers.

These tables have been prepared by taking off from the brother-brother data papers of my school measurement records all the available pairs of cases falling into each series. Thus in some cases the ages of both brothers were given, but not the head measurement of one or other; in other cases the head measurements of both, but the age of one or other would fail, or again the age of one and the head measurement of the other might be all the information available. Thus the total number of cases and the frequency distribution varies slightly from one table to a second.

* The measurements form part of the material obtained with the assistance of a grant from the Royal Society Government Grant Committee.

A few remarks must be made on these tables.

Table I gives the following values of the constants :—

Mean length of head of elder brother	=	186·7508 in mm.
" " younger "	=	183·8296 "
Standard deviation of elder brother	=	7·5027 "
" " younger "	=	7·3536 "

The correlation is, then, found to be 0·601,751,* and the regression, younger on elder brother, 0·5897. These give the intensity of heredity, uncorrected, for the growth factor.

Now, the most noteworthy part of this result is, as we shall see later, that *taking brothers at different ages tends to exaggerate the apparent intensity of heredity*. If we were to take pairs of boys at ages from 4 to 19, each pair having no hereditary relationship, but being, on the average, within a year or so of the same age, we should find a spurious correlation due to the mixture of material, each pair having approximately-like head-lengths because the members of it were, approximately, of like age. On the other hand, if the boys were blood relations of very different ages, their apparent relationship would be weakened, because we should be correlating the same organ at different stages of its growth. We have thus two factors : one tending to exaggerate, and the other to weaken the apparent strength of hereditary resemblance. It is of great interest to note that the former factor in the present case is the more effective.

In Table II we have what I term a growth table, *i.e.*, a correlation table between period of growth and the quantitative measure of a character. The constants of this table are as follows :—

Mean age of boy	=	13·0394 years.
Standard deviation of age............	=	2·8207 "
Mean head-length	=	185·4516 mm.
Standard deviation of head-length	=	7·4991 "
Correlation of age and head-length	=	0·453,496

The regression coefficient for head-length on age = 1·205676, and we have the probable head-length H_P for observed age A given by

$$H_P = 169 \cdot 7303 + 1 \cdot 2057\,A \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\epsilon.)$$

Thus, on the average, boys' heads grow in length 1·2 mm. a year.

My results are based on 1637 cases entirely taken off the brother-brother data papers. Dr. Alice Lee at an earlier stage also worked out a growth table. We had not then so many brother-brother data papers filled in. She used in addition all the brother measurements on the brother-sister papers, and so reached 1856 boys, of which, I

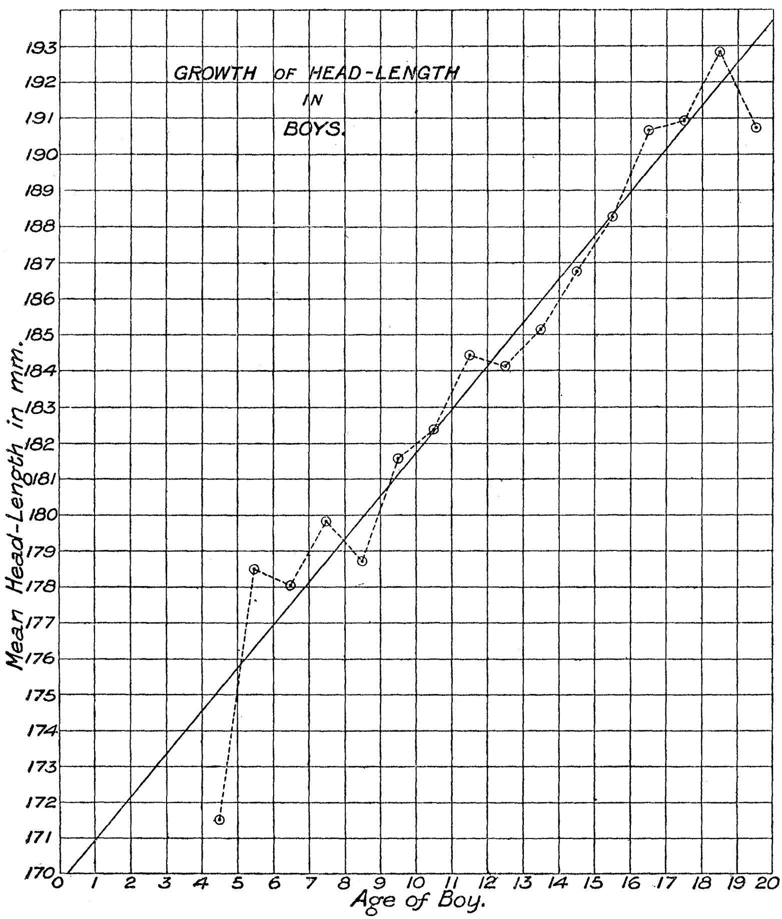
* Six figures have been kept in the correlation coefficients, as we require to calculate the regression coefficients from the differences of products and powers.

think, we may safely assert that 400 at least are not included in my series. She found: Mean age,* 12·7177; mean head-length, 184·8182; and slope of regression line, 1·2040, giving the formula

$$H_P = 169 \cdot 5061 + 1 \cdot 2040\,A,$$

a result in substantial agreement with mine.

DIAGRAM 1.



In Diagram 1 the formula (ϵ) is represented with the observed mean values at each year of life. The results for the 4th, 5th, and 19th years of life ought not to be considered, for they are based on only 2, 10, and 12 observations respectively. It will clearly hardly be possible to express the growth curve better than by a straight

* The mean age is less, because brother-sisters are obtained chiefly from primary, not secondary, schools.

line, until the range of data is very largely extended. The regression is sensibly linear.

Table IIIA and Table IIIB give the following results :—

Mean age of elder brother ..	= 14·1249	Mean age of younger brother	= 11·7149
S.D. of elder brother's age. .	= 2·5124	S.D. of younger brother's age	= 2·7221
Mean head-length, younger brother................	=183·8578	Mean head-length, elder brother................	=186·6515
S.D. of head-length, younger brother	= 7·2806	S.D. of head-length, elder brother................	= 7·5005
Correlation of age of elder and head-length of younger ..	= 0·396,598	Correlation of age of younger and head-length of elder .	= 0·379,326

We see accordingly that within the limits of the probable error, the correlation between younger brother's head-length and elder brother's age is the same as that between elder brother's head-length and younger brother's age. This result might, to some extent, have been anticipated, but actual proof of this type of cross-relation is of value. In Table IV we have the correlation between ages of brothers giving the constants :—

Mean age of elder brother	= 14·1508
Mean age of younger brother	= 11·7487
S.D. of elder brother's age	= 2·5080
S.D. of younger brother's age	= 2·7220
Correlation of brothers' ages	= 0·884,186

The first four results are in good agreement with those of Tables IIIA and IIIB. The last result shows how nearly there is an approximation to a constant difference in age between brothers in schools. Very closely we have—

Probable age of younger brother $= 0 \cdot 96 \times$ (age of elder brother) $- 1 \cdot 83$.

When the elder brother is 6, his younger brother is probably 2·1 years younger than he is; when the elder brother is 12, the younger brother is probably 2·3 years younger, and when he is 18, 2·6 years younger. The explanation of this is that when the elder brother is very young only his near or second brother will, as a rule, be at the same school, but in the secondary schools, which he reaches at a much later age, it is possible for a much younger brother to be at the same school.

Now let us substitute the correlation values, found in equations (i) to (iii), of page 290. We have

$$r_{xy} = 0 \cdot 601,654, \qquad r_{t_1 t_2} = 0 \cdot 884,186$$
$$r_{xt_1} = r_{yt_2} = 0 \cdot 453,496,$$
$$r_{xt_2} = 0 \cdot 379,326, \qquad r_{yt_1} = 0 \cdot 396,598.$$

Whence we find

$$\Sigma_x/\sigma_x = 0{\cdot}890{,}051, \qquad \Sigma_y/\sigma_y = 0{\cdot}891{,}209,$$

and

$$R = 0{\cdot}5446.$$

This is a very reasonable value of fraternal correlation, agreeing quite well with results obtained for horse, man and dog. It is worth noting that

$$r_{xt_1} \times r_{t_1t_2} = r_{yt_2} \times r_{t_1t_2} = 0{\cdot}4010,$$

and, therefore, either equals r_{xt_2} or r_{yt_1} fairly closely; in fact, within the probable error of their difference.

Hence, it would appear highly probable that the cross-relation between one brother's head length and a second brother's age is solely due to the correlation of the ages between the two brothers.

If such a result as

$$r_{xt_1} \times r_{t_1t_2} = r_{xt_2} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (\eta)$$

should be verified on the reduction of further data, it will enable us to much simplify our formulæ.

Thus we easily find for this case

$$\Sigma_x = \sigma_x\sqrt{1 - r_{xt_1}{}^2}, \qquad \Sigma_y = \sigma_y\sqrt{1 - r_{yt_2}{}^2}$$

and

$$R = \frac{r_{xy} - r_{xt_1}{}^2 r_{t_1t_2}}{1 - r_{xt_1}{}^2}$$

Or, we require to find only the uncorrected correlation (ρ) the growth correlation (r), and the correlation between periods of growth ($\mathbf{r}$). The correction to be made to the apparent correlation is then the subtraction from it of

$$\frac{r^2(\mathbf{r} - \rho)}{1 - r^2}$$

I hope shortly to ascertain whether relations like η above hold also for other head-measurements on growing children.

Table I.—Collateral Heredity Head Length in Brothers, uncorrected for Age.

*Head-length of Elder Brother in mm.**

Head-length of Younger Brother in mm. 162—3.	162—3.	164—5.	163—7.	168—9.	170—1.	172—3.	174—5.	176—7.	178—9.	180—1.	182—3.	184—5.	186—7.	188—9.	190—1.	192—3.	194—5.	196—7.	198—9.	200—1.	202—3.	204—5.	206—7.	208—9.	Totals.
162—3	—	—	2	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	3
164—5	—	—	—	—	—	—	—	—	1	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	2
166—7	—	—	—	—	1	1	0·5	1·5	—	—	—	—	—	—	—	—	0·25	0·25	—	—	—	—	—	—	4·5
168—9	—	—	1	0·5	1·5	1	1	1	2	1	1	1	1	1	—	—	0·25	0·25	—	—	—	—	—	—	13·5
170—1	1	—	—	4	3	3	—	3	4	2	3	3·5	3	3	—	—	1	—	—	—	—	—	—	—	33·5
172—3	—	—	—	1·75	4·25	2	2	3	1	2	2	2·5	1	1	—	1	—	—	—	—	—	1	—	—	24·5
174—5	—	—	1	0·75	3·25	1	—	5	5	9	3·5	8	2	5	—	1	1	—	—	—	—	—	—	—	45·5
176—7	—	—	—	—	1	1	6	3	8	5·5	10·5	10	4	3	2	—	1	—	1·5	—	—	—	—	—	56·5
178—9	—	—	—	—	2	3	2·5	6·5	11	8	5	11·5	8	5·5	4·5	2	3	1	1·5	—	—	—	—	—	75
180—1	—	—	—	5	1·5	—	2	4	11	17	13·5	14·5	16	14	6·5	8·5	6	2·5	—	—	1	1	—	—	119·5
182—3	—	—	—	—	2	2	2·5	4·5	8	9	13	14·5	8·5	12	11	10	2	1·75	1·25	1	—	—	—	—	103
184—5	—	—	—	—	—	—	—	2	1	11·5	12	17	15·5	17·75	9·75	8·25	7·25	4·25	3·25	2	2	—	—	—	113·5
186—7	—	—	—	—	—	1	2	2	2	6	12·5	7	8	15·25	17·75	11·75	8·25	5	1	—	—	—	—	—	99·5
188—9	—	—	—	1	—	1	1·5	1·5	2	4	5	7	6	9·5	12	6·5	7·5	4·5	1	1	1	—	—	—	72
190—1	—	—	—	—	—	—	—	—	—	2	1	3·5	8·5	10·5	9·5	7·5	10·5	10	3	2	—	2	—	—	70
192—3	—	—	—	—	—	—	—	—	2	1	2	2	4	7	10·5	7·5	8	6	7	1·5	—	1	—	1	60·5
194—5	—	—	—	—	—	—	—	—	—	1	1	1	1	5·5	4	3	2·5	5	5	3·5	1	1	—	—	34·5
196—7	—	—	—	—	—	—	—	—	—	—	—	1	1	1	3·5	2	5	0·5	3·5	—	—	1	—	—	18·5
198—9	—	—	—	—	—	—	—	—	—	—	—	—	2	—	—	2	2	4	2	4	2	1	—	—	19
200—1	—	—	—	—	—	—	—	—	—	—	—	—	1	—	1	—	—	1	1	2	1	1	1	—	9
202—3	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	1	2	—	—	4
204—5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1 5	0·5	—	—	1	—	—	—	—	—	3
206—7	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	1
208—9	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	1
Totals	1	0	4	9·5	19·5	16	20	37	58	79	85	104	92·5	112	93·5	71·5	65·5	47	32	17	9	11	1	1	986

* The group 162—3 contains all boys with head-lengths greater than 161·5 and less than 163·5 and so on. When a head-length fell exactly on such a value as 183·5 it was halved between the columns 182—3 and 183—4.

Table II.—Correlation of Age and Length of Head in Boys.

*Year of Age of Boy.**

Length of Head in mm.†	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	Totals.
162—3	—	—	—	—	1	1	—	—	—	—	—	1	—	—	—	—	3
164—5	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	2
166—7	—	—	1	0·5	2	1	1	—	1	—	—	—	—	—	—	—	6·5
168—9	—	—	2	1·5	2	4	2	2	1	0·5	—	1	—	—	—	—	16
170—1	1	1	2	2·5	3	4	4	5·5	6	5·5	3	1	—	1	—	—	39·5
172—3	1	—	—	1·5	7	4·5	3	2·5	11	5	1	2	—	—	—	—	38·5
174—5	—	1	0·5	4	6	5·5	9	4	8	2·5	3·5	3	1	1	—	—	49
176—7	—	—	0·5	10	4	10	12	6	13	12	4·5	4	1	—	—	—	77
178—9	—	5	2	2	10·5	10	10	18	14	11·5	9	5	5	1	—	—	103
180—1	—	—	2	5	7·5	22	23	17	17	27	18	9·5	8	5	—	—	161
182—3	—	3	1	4	9	6	15	22	26·5	25	17	16	6	2	3	2	157·5
184—5	—	—	1	4	3·5	10·5	19	18	31	36	25	23	13·5	6	1	1	192·5
186—7	—	—	1	5	6·5	13	24	16·5	24·5	24	17·5	13·5	3·5	8	1	—	158
188—9	—	—	1	1	3	5·5	16	15	23	20	22·5	17	19·5	9	9	1	162·5
190—1	—	—	1	—	—	7	6	14·5	22	18·5	19·5	19·5	16	9	4	0·5	137·5
192—3	—	—	—	2	—	5	2	10	12	18·5	17·5	18	10·5	6·5	6	4·5	112·5
194—5	—	—	—	1	—	1	—	7	11	8·5	3	18·5	15·5	6·5	4	1	77
196—7	—	—	—	—	1	—	2	3	3	7·5	11	7	9	4	6	1	54·5
198—9	—	—	—	—	—	—	—	3	3	3	7	8	4·5	12	1	1	42·5
200—1	—	—	—	—	—	1	—	2	—	—	2	5	5·5	3	1	—	19·5
202—3	—	—	—	—	—	—	—	—	1	1	1	1	3·5	1	1	—	9·5
204—5	—	—	—	—	—	—	—	—	—	1	—	2	6	2	1	—	12
206—7	—	—	—	—	—	—	—	1	—	—	—	—	1	—	1	—	3
208—9	—	—	—	—	—	1	—	—	—	—	—	1	—	1	—	—	3
Totals..	2	10	15	44	66	112	148	168	229	227	182	176	129	78	39	12	1637

* All boys from the nth to the $(n+1)$th anniversary of their birthday would be placed in the nth column, or the mean age of such boys $= n + 0{\cdot}5$ years.

† See Note to Table I.

Table IIIA.—Correlation of Age of Elder and Head-length of Younger Brother.

*Year of Age of Elder Brother.**

Head-length of Younger Brother in mm.†	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	Totals.
162—3	—	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	2
164—5	—	—	—	—	—	—	—	—	1	—	1	—	—	—	—	—	2
166—7	—	—	—	—	1	—	1	1	1	0·5	—	—	—	—	—	—	4·5
168—9	—	—	—	—	—	2	1	6	2	2·5	1	—	—	—	—	—	14·5
170—1	—	1	1	—	2	1·5	2	5·5	7·5	6	1·5	3	1	—	—	—	32
172—3	—	—	—	—	2	0·5	4·5	3·5	6	2·5	2·5	2	2	—	—	1	26·5
174—5	—	—	—	—	1·5	3	4·5	3	10	7·5	6·5	2	3	1	—	—	42
176—7	—	—	—	—	1·5	2	8	8	6·5	10	4·5	8	2	2·5	—	—	53
178—9	—	—	2	—	1	2	10	9	15·5	11	8	5	8	2·5	—	—	74
180—1	—	—	—	—	1	5	14	12	20·5	16	16	14	9·5	7·5	5	2	122·5
182—3	—	1	—	—	2	4	5	7	14	19	8	17	12	8·5	1	3	101·5
184—5	—	—	—	—	—	—	5·5	5	16·5	15·5	18	21	17·5	9·5	5	2	115·5
186—7	—	—	—	—	—	2	4·5	9	20	19	10	15	8	3·5	7	—	98
188—9	—	—	—	—	1	3	2	7	8	16	7·5	5·5	10	10·5	5	1	76·5
190—1	—	—	—	—	—	—	—	4	6·5	10·5	12·5	9·5	7	11·5	5	1	67·5
192—3	—	—	—	—	—	—	3	1	3	10	8	9	13	7·5	5	3	62·5
194—5	—	—	—	—	—	—	1	—	4	3	5	3	5	3·5	5	4	33·5
196—7	—	—	—	—	—	—	—	—	3	2	2	6	4	2	—	—	19
198—9	—	—	—	—	—	—	—	—	—	2	2	3	6	3	1	1	18
200—1	—	—	—	—	—	—	—	1	—	1	—	—	1	2	2	—	7
202—3	—	—	—	—	—	—	—	—	—	—	—	—	1	1	1	—	3
204—5	—	—	—	—	—	—	—	—	—	—	—	1	1	1	—	1	4
206—7	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1
208—9	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	—	1
Totals..	—	2	3	—	13	25	67	83	146	154	115	124	111	77	42	19	981

* See Note to Table II.

† See Note to Table I.

Table IIIB.—Correlation of Head Length of Elder and Age of Younger Brother.

*Year of Age of Younger Brother.**

Head-length of Elder Brother in mm.†	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	Totals.
162—3	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	1
164—5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
166—7	—	1	—	—	—	2	—	—	—	1	—	—	—	—	—	—	4
168—9	—	—	1	2	1	2·5	0·5	—	1·5	—	—	—	—	—	—	—	8·5
170—1	—	1	—	1	4	2·5	2·5	2	3·5	1	—	1	—	—	—	—	18·5
172—3	—	1	—	2	2	3	2	2	2	2	—	—	—	—	—	—	16
174—5	—	—	2	3	2	3·5	2	—	3·5	2	1	1	—	—	—	—	20
176—7	—	1	2	5	4	8·5	3	5	4·5	1	2	1	—	—	—	—	37
178—9	1	2	2	4	8	10	6	8	6	4	3	3	—	—	—	—	57
180—1	—	2	2	4	5	9·5	6·5	16	9	8	6	3	—	—	—	—	71
182—3	—	1	2	5	9	14·5	14	11	8	14·5	5	2	1	2	—	—	89
184—5	—	1	1	10	12	12·5	14	12	20·5	6·5	10	4·5	4	—	—	—	108
186—7	—	—	3	9	8	14·5	18	8·5	12	10	4	3·5	2	—	—	—	92·5
188—9	2	1	2	4	6	12·5	11·5	18	12·5	16·5	7	11	8	—	—	—	112
190—1	—	2	2	4	3	9·5	11	11	10	12·5	12·5	9	1	1·5	—	—	89
192—3	—	1	1	4	1	3	7	10·5	7	11	6·5	11·5	5·5	1·5	1	—	71·5
194—5	—	—	—	3·5	2	4	8	7	6·5	9	12	9·5	4·5	2	—	—	68
196—7	—	—	—	0·5	1	2	1	5·5	5·5	9·5	7	10	2	1	—	—	45
198—9	—	—	—	1	—	2	1	3·5	2	2·5	9	7	3	1	—	—	32
200—1	—	—	—	—	—	1	1	0·5	1	2	2	6·5	—	—	—	—	14
202—3	—	—	—	—	—	1	—	0·5	1	1	2	1·5	1	—	—	—	8
204—5	—	—	—	—	—	—	—	1	3	1	6	—	1	—	—	—	12
206—7	—	—	—	—	—	—	—	—	—	—	2	—	—	—	—	—	2
208—9	—	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	1
Totals	3	14	20	62	68	118	109	122	119	116	97	86	33	9	1	—	977

* See Note to Table II.

† See Note to Table I.

Table IV.—Correlation of Ages of Pairs of Brothers.

*Years of Age of Elder Brother.**

Year of Age of Younger Brother.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	Totals.
4	—	1	—	—	1	—	1	1	—	—	—	—	—	—	—	—	4
5	—	1	2	—	3	—	3	2	2	—	—	—	—	—	—	—	13
6	—	—	1	—	4	4	5	5	1	—	—	—	—	—	—	—	20
7	—	—	—	—	4	10	10	14	8	8	2	2	—	—	—	—	58
8	—	—	—	—	—	10	26	13	16	6	1	—	—	—	—	—	72
9	—	—	—	—	—	1	17	35	32	16	7	2	1	1	—	—	112
10	—	—	—	—	—	—	1	17	55	26	8	3	2	—	—	—	112
11	—	—	—	—	—	—	—	—	28	58	25	11	7	—	—	—	129
12	—	—	—	—	—	—	—	—	3	41	39	16	11	6	—	1	117
13	—	—	—	—	—	—	—	—	—	—	30	60	14	6	2	2	114
14	—	—	—	—	—	—	—	—	—	—	5	27	43	18	7	5	105
15	—	—	—	—	—	—	—	—	—	—	—	2	32	32	12	5	83
16	—	—	—	—	—	—	—	—	—	—	—	—	1	17	14	2	34
17	—	—	—	—	—	—	—	—	—	—	—	—	—	1	5	5	11
18	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	1
19	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
Totals	—	2	3	—	12	25	63	87	145	155	117	123	111	81	40	21	985

* See Note to Table II.

This table contains only fifteen pairs of twins, one pair of brothers having been born within ten months of each other. Any age, as 9, means all falling within the ninth year, *i.e.*, from the ninth birthday to the day before the tenth birthday, so that two brothers, not twins, might appear of the same age.

(4.) *To find the correction to be made to the apparent homotypic correlation, when each pair of homotypes is differentiated by a common period of growth from other pairs of homotypes.*

The solution of this problem may be deduced at once from equation (iii) of the preceding problem by simply putting $t_1 = t_2$. In this case $r = r'$, $\mathbf{r} = 1$, and we find

$$\mathrm{R} = \frac{\rho - r^2}{1 - r^2} \quad \text{(v).}$$

This equation was given by me in a note in *Biometrika*, vol. 1, p. 404, and its use illustrated on Dr. Simpson's data for *Paramecium caudatum.*

(5.) *To find the correction to be made to the apparent homotypic correlation when the pair of homologous parts are differentiated from each other by situation on the organism.*

We have only to put in formula (iii) on p. 290, $t_1 = p_1$ and $t_2 = p_2$, the positional co-ordinates of the first and second homologous parts, to make that formula available for position instead of age differentiation. If we denote by c_1 and c_2 the characters of the parts in the positions p_1 and p_2 respectively, our solution takes the form below, where we have confined our attention to the same character,

$$\mathrm{R} = \frac{\rho(1 - r_{p_1p_2}{}^2)}{1 - r_{p_1p_2}{}^2 - r_{p_1c_1}{}^2 - r_{p_1c_2}{}^2 + 2r_{p_1p_2}r_{p_1c_1}r_{p_1c_2}}$$
$$- \frac{2r_{p_1c_1}r_{p_1c_2} - r_{p_1p_2}(r_{p_1c_1}{}^2 + r_{p_1c_2}{}^2)}{1 - r_{p_1p_2}{}^2 - r_{p_1e_1}{}^2 - r_{p_1c_2}{}^2 + 2r_{p_1p_2}r_{p_1c_1}r_{p_2e_2}} \ldots \text{(vi).}$$

This follows since $r_{p_1c_1} = r_{p_2c_2}$, and $r_{p_1c_2} = r_{p_2c_1}$.
We have again, therefore, to find four correlation coefficients. But this formula simplifies immensely if we observe the following conditions :

(*a*.) Take the same number of homotypes or homologous parts from the same positions in each organism.

(*b*.) Enter each one of these homotypes or homologous parts with each other on the same organism, so as to obtain a symmetrical table, *i.e.*, c_1 is first entered with c_2 and then c_2 with c_1,

These conditions are or can be usually satisfied in any homotyposis investigation.

(6.) Further, the positions will, as a rule, be arranged in series and may be numbered $1, 2, 3, 4, \ldots m$, if m homotypes or homologous parts be taken from each individual organism. The position scale is, of course, perfectly arbitrary, and has nothing to do, for example, with the actual distances between positions on the organism. We can make it a uniform numerical scale, which for convenience we can take to be the same serial order as that of positions on the organism.

Let $\bar{p}$ = mean position, σ_p = standard deviation of positions on the arbitrary position scale. Let there be n organisms, and suppose that S_m denotes a summation of all m homologous parts on an organism, and S a summation for all n organisms. Then, if $\sigma_p = \sigma_{p_1} = \sigma_{p_2}$,

$$nm(m-1)r_{p_1p_2}\sigma_p^2 = S_n\{S_m(p_1-\bar{p})(p_2-\bar{p})\}$$
$$= S_n\{S_m(p_1-\bar{p}) \times S_m(p_2-\bar{p}) - S_m(p_1-\bar{p})^2\}$$

But $S_m(p_1-\bar{p}) = 0$, hence, since $S_m(p_1-\bar{p})^2 = m\sigma_p^2$,

$$nm(m-1)r_{p_1p_2}\sigma_p^2 = -nm\sigma_p^2,$$

or

$$r_{p_1p_2} = -\frac{1}{m-1} \quad \text{(vii).}$$

Further

$$nm(m-1)r_{p_1c_2}\sigma_p\sigma_c = S_n\{S_m(p_1-\bar{p})(c_2-\bar{c})\}$$
$$= S_n\{S_m(c_2-\bar{c}) \times S_m(p_1-\bar{p}) - S_m(c_1-\bar{c})(p_1-\bar{p})\}$$
$$= -S_n\{S_m(c_1-\bar{c})(p_1-\bar{p})\}.$$

But $S_n\{S_m(c_1-\bar{c})(p_1-\bar{p})\} = nm\,\sigma_c\sigma_p r_{c_1p_1}$.

Hence

$$r_{p_1c_2} = -\frac{r_{c_1p_1}}{m-1} = r_{c_1p_1} \times r_{p_1p_2} \quad \text{(viii),}$$

a relation precisely similar to that discovered in the case of growth periods for brother's head-lengths from the actual numbers on p. 294.

Substituting we find

$$1 - r_{p_1p_2}^2 - r_{p_1c_1}^2 - r_{p_1c_2}^2 + 2r_{p_1p_2}r_{p_1c_1}r_{p_1c_2} = (1-r_{p_1p_2}^2)(1-r_{p_1c_1}^2)$$
$$2r_{p_1c_1}r_{p_1c_2} - r_{p_1p_2}(r_{p_1c_1}^2 + r_{p_1c_2}) = r_{p_1p_2}r_{p_1c_1}^2(1-r_{p_1p_2}^2)$$

Then substituting in (vi) and using (vii) we determine the simple formula for homotyposis corrected for positional differentation

$$R = \frac{\rho}{1-r_{pc}^2} + \frac{1}{m-1}\,\frac{r_{pc}^2}{1-r_{pc}^2} \quad \text{(ix).}$$

where r_{pc} stands for the correlation of character and position on the organism.

An exactly similar formula might be found for the correction for the age or growth factor, if the m homologous parts dealt with had the same distribution of ages or growths in each organism.

(7.) Now the equation just found has the serious disadvantage that it is based on the linearity* of the regression relation between position

* The reader should note that this condition does not involve any assumption of normal frequency, or the Gaussian law. The latter applies only to a very special case of linear regression.

and character. But while organic and homotypic correlations give for a surprising variety of cases sensibly linear regression relations, the relation between position and mean character is far more rarely linear. We obtain, as a rule, remarkably smooth curves. We, therefore, require some modification of equation (ix).

Still supposing the regression of character and position linear, we should have, if σ' be the mean standard deviation of an array of organs in the same position,

$$\sigma'^2 = \sigma^2(1 - r_{pc}^2).$$

But if σ_M be the standard deviation of the means of the arrays, we have from first principles

$$\sigma^2 = \sigma_M^2 + \sigma'^2$$

Hence $$\sigma_M = \sigma \times r_{pc}.$$

We can now write equation (ix) in the form

$$R = \rho\frac{\sigma^2}{\sigma^2 - \sigma_M^2} + \frac{\sigma_M^2}{(m-1)(\sigma^2 - \sigma_M^2)} \quad \ldots\ldots\ldots\ldots\ldots \quad \text{(x)}.$$

This is quite free from r_{pc}, and, what is more, although we have deduced it from (ix) and the relation $\sigma'^2 = \sigma^2(1 - r_{pc}^2)$ peculiar to linear regression, it is now free of any limitation as to the nature of the relation between position and mean character. Thus (x) is a far more important formula than (ix), and should always be used, until we have shown that the relation between position and mean character is sensibly linear. If anything, it involves less arithmetic than (ix).

We can show this *ab initio* as follows:—Let the individuality of the organism in any homologous part be measured by its excess above (respectively defect below) the mean value of the character for the homologous part in that position.* Then, if c' = element of character due to individuality, and $\bar{c}_p$ be the mean character in any position for the n individuals dealt with,

$$c_1' = c_1 - \bar{c}_p, \quad S_n(c_1) = n\bar{c}_p, \quad \text{and} \quad S_n(c_1') = 0.$$

Hence we easily find

$$S_n(c_1'^2) = S_n(c_1^2) - n\bar{c}_p^2$$

$$S_mS_n(c_1'^2) = S_mS_n(c_1^2) - S_m(n\bar{c}_p^2),$$

* It might be considered better, if the standard deviations of the homologous parts vary very considerably with position, to measure the individuality by the ratio of this excess to the corresponding standard deviation. Not only, however, does the use of such a ratio immensely increase the arithmetical labour, which is a possibility, which of course, we could face, but there is also a question as to whether the ratio is really a *truer* measure of individuality. A full discussion of this important point must for the present be deferred.

Or, noting that

$$\bar{c} = S_m S_n(c_1)/(nm) = S_m(\bar{c}_p)/m, \text{ we have}$$

$$\sigma'^2 = \sigma^2 - \sigma_M{}^2,$$

where σ' is the standard deviation of the character-individualities free from the position factor. We see that it is precisely the same quantity as we have previously used for the mean standard deviation of the arrays for given positions.

Next taking the correlation of characters c'_1 and c'_2 in positions p_1 and p_2 we have

$$S_m(c_1'^2) + S_m(c_1'c_2') = S_m(c_1^2) + S_m(c_1c_2) - 2S_m(c_1)m\bar{c} + m^2\bar{c}^2.$$

To get this result we have multiplied every quantity like $c'_1 = c_1 - \bar{c}_{p_1}$ by every other quantity like $c'_2 = c_2 - \bar{c}_{p_2}$ *and by itself*, and then added such quantities together for every position on the one organism. Thus on the left hand side there are m terms in the first, $m(m-1)$ terms in the second summation; on the right hand side there are m terms in the first, $m(m-1)$ terms in the second and m terms in the third summation. Now sum for each of the n organisms, and we have

$$nm\sigma'^2 + nm(m-1)R\sigma'^2 = mn(\sigma^2 + \bar{c}^2) + nm(m-1)(\rho\sigma^2 + \bar{c}^2) - 2m^2n\bar{c}^2 + m^2n\bar{c}^2.$$

Whence

$$R = \rho\sigma^2/\sigma'^2 + \frac{\sigma^2 - \sigma'^2}{(m-1)\sigma'^2},$$

or, as before

$$R = \rho\frac{\sigma^2}{\sigma^2 - \sigma_M{}^2} + \frac{\sigma_M{}^2}{(m-1)(\sigma^2 - \sigma_M{}^2)} \quad \text{.............. (x).}$$

Now while this proof is independent of the theory of partial correlation coefficients, involving only simple algebra, and is further independent of any consideration of linear regression, it yet wants something of the width of the former theory, which allows us at once, for example, to correct for a combination of factors, such for example as for *both* growth and position influences simultaneously. The difficulty lies entirely in the extent within which it is legitimate to assume the relation between position or age, and the mean value of the character at that position or age to be linear. It is therefore clearly advisable to start by plotting this relationship,* and fitting, if possible, such position or growth graphs with appropriate curves. If, for the series of positions dealt with or the period of growth taken, we find that a straight line† is a close approximation to the relationship, then we

* In the case of some animals and many plants the relationship is in itself of much interest, for it expresses a law of development or growth in serial parts.

† The analytical consideration of this point is very simple. If the regression

may use the general theory of partial correlation, otherwise we must fall back on results like (x). For example, in head growth in boys, we cannot much improve on a straight line; in positional influence on the branches in the whorls of *Equisetum arvense* we need at least a third order parabola.

(8.) Although material for several investigations on the homotyposis of serial homologous parts has been collected, the progress in some of these cases is slow, as it involves rather laborious microscopic measurement. I content myself at present with an illustration from the vegetable kingdom.

I collected in the autumn of last year, 126 plants of *Equisetum arvense* in Raydale Side, an offshoot of Wensley Dale; the plant was growing on a lane side high up above Semmerwater. This *Equisetum* grows from the top with a single stem, and I counted the number of branches to the whorl from the root upwards. As a rule, there will be one or two whorls close to the soil which have never developed any branches at all; then we have what I shall term the *first* whorl in which some branches have developed, but the number is irregular and obviously subject to some cause of variation, other than the growth law of the plant. The number of branches to the whorl then increases uniformly and steadily up to the 4th whorl, after which it falls almost equally steadily to the 10th whorl. Beyond this the results becomes somewhat irregular again, for very few plants will be found—at any rate in the locality considered—with more than 12 or 13 whorls, and even in these whorls there is a certain amount of forking or irregularity which it is difficult to deal with. The plants were certainly fully developed

be linear, the means of the arrays all lie on the regression line, and the mean standard deviation of the arrays about their means is $\sigma\sqrt{1-r^2}$. If the regression be not linear, the means of the arrays will have a mean square deviation $\Sigma_M{}^2$ from the regression line. The mean square deviation of the arrays from the regression line, but *not from their means*, is still $\sigma^2(1-r^2)$. The mean standard deviation (deviation of mean square *from means*) is now given by

$$\sigma'^2 = \sigma^2(1-r^2)-(\sigma_M{}^2-r^2\sigma^2),$$

since $\sigma^2 = \sigma'^2+\sigma_M{}^2$. But we easily find

$$\Sigma_M{}^2 = \sigma_M{}^2-r^2\sigma^2.$$

Hence Σ_M is a good measurement of the deviation of regression from linearity, or of σ_M from $r\sigma$. If we take $\eta = \sigma_M/\sigma$, we have

$$\sigma'^2 = \sigma^2(1-\eta^2), \qquad \Sigma_M{}^2 = (\eta^2-r^2)\sigma^2.$$

Clearly η^2 must lie between r^2 and 1. Further, η can only vanish when the correlation is zero, or become ± 1 when the correlation is perfect. Between these values it gives the mean reduction in variability of an array as compared with the whole population. Further, the deviation of η from r is a good measure of the deviation of the system from linearity. Thus η is a useful constant which ought always to be given for non-linear systems. It measures the approach of the system not only to linearity but to a single valued relationship, *i.e.*, to a causal nexus.

when gathered at least as far as the 12th or 13th whorl, and I doubt whether even beyond this so late in the season, any further branching would have taken place. A few branches were broken off, and these were of course counted; there was no difficulty, however, in easily ascertaining whether a branch had in any case been developed or not, and the peculiarity of the 1st whorl was certainly not due to missing, but to undeveloped branches.

Table V gives the relation between branches to the whorl and position for the whole of the 126 plants. In two columns to the right are given the means and variabilities of the branches for each whorl.

Now, whether we judge by mean or standard deviation, we see a perfectly gradual change from whorl to whorl, which absolutely precludes us from considering the number of branches to the whorl as a pure homotypic character. We see a marked differentiation due to position of the whorl on the plant; the whorls are homologous but not homotypic parts.

Suppose, however, that we disregard our test for differentiation,* and proceed to find a correlation table for the whole material as homotypic. We have Table VI, for which I have to heartily thank Dr. Alice Lee.

The value found for the homotypic correlation from this table is

$$\rho = -0 \cdot 0064 \pm 0 \cdot 0185,$$

or, there is no sensible homotyposis at all.

But we might have gone to the other extreme and taken only the 3rd, 4th, and 5th whorls, which have more nearly the same means and standard deviations as homotypes. The result is Table VII, giving

$$\rho = 0 \cdot 7918 \pm 0 \cdot 0129.$$

It will be perfectly clear, therefore, as these two results ought to be the same, if the whorls were true homotypes, that we may get any result at all if we neglect differentiation.† The answer to this is that no trained biometrician would call these whorls "undifferentiated like organs" with the two right hand columns of Table V before him.

* On the test for differentiation, see 'Biometrika,' vol. 1, p. 334.

† Bateson, 'Roy. Soc. Proc.,' vol. 69, p. 200.

Table V.

Number of Branches to the Whorl.

Position of Whorl.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	Totals.	Mean.	Standard deviation.
1	2	2	3	10	9	8	13	29	22	17	11	—	—	126	7·619	2·360
2	—	—	—	—	1	3	5	21	37	40	16	3	—	126	9·294	1·273
3	—	—	—	—	—	—	9	9	35	45	23	5	—	126	9·627	1·187
4	—	—	—	—	—	—	6	10	33	45	28	3	1	126	9·730	1·151
5	—	—	—	—	—	—	8	10	35	41	30	2	—	126	9·643	1·158
6	—	—	—	—	2	3	6	13	35	38	24	3	—	124	9·427	1·363
7	—	—	1	4	2	6	12	23	28	29	17	1	—	123	8·732	1·781
8	2	3	7	5	5	13	21	23	24	14	4	—	—	121	7·297	2·291
9	8	10	13	14	9	14	19	17	10	5	—	—	—	119	5·555	2·553
10	18	20	13	11	17	14	11	5	1	—	—	—	—	110	3·964	2·199
11	31	29	18	9	5	3	1	1	—	—	—	—	—	97	2·443	1·506
12	24	34	6	2	—	—	1	—	—	—	—	—	—	67	1·866	0·960
13	24	14	—	—	1	—	—	—	—	—	—	—	—	39	1·462	0·746
14	8	4	—	—	—	—	—	—	—	—	—	—	—	12	1·333	0·471
15	3	1	—	—	—	—	—	—	—	—	—	—	—	4	1·250	0·433
16	2	—	—	—	—	—	—	—	—	—	—	—	—	2	1·000	0·000
Totals . . .	122	117	61	55	51	64	112	161	260	274	153	17	1	1448	—	—

Table VI.

Number of Branches on 1st Whorl of Pair. (columns)

Number of Branches on 2nd Whorl of Pair. (rows)

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	Totals.
1	132	110	63	55	41	61	91	120	281	274	146	22	3	1399
2	110	98	49	49	57	54	90	129	281	237	160	18	1	1333
3	63	49	16	18	14	31	59	79	122	149	55	6	—	661
4	55	49	18	12	19	24	53	82	103	121	55	1	1	593
5	41	57	14	19	18	22	67	45	115	109	58	5	—	570
6	61	54	31	24	22	8	54	70	133	154	54	7	—	672
7	91	90	59	53	67	54	162	122	199	173	72	4	2	1148
8	120	129	79	82	45	70	122	300	281	262	111	28	2	1631
9	281	281	122	103	115	133	199	281	766	401	150	9	1	2842
10	274	237	149	121	109	154	173	262	401	914	240	25	2	3061
11	146	160	55	55	58	54	72	111	150	240	428	40	1	1570
12	22	18	6	1	5	7	4	28	9	25	40	34	—	199
13	3	1	—	1	—	—	2	2	1	2	1	—	—	13
Totals	1399	1333	661	593	570	672	1148	1631	2842	3061	1570	199	13	15692

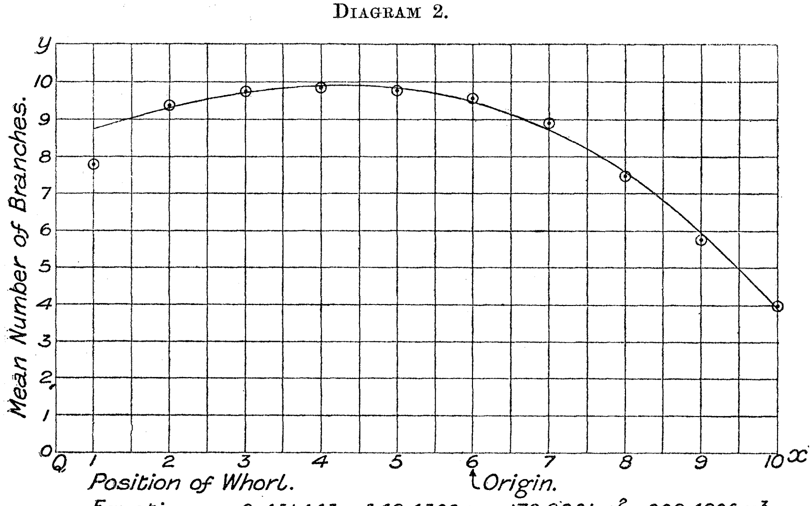
Table VII.—Whorls, 3rd, 4th, and 5th only.

Number of Branches to 1st Whorl of Pair.

Number of Branches on 2nd Whorl of Pair.	7.	8.	9.	10.	11.	12.	13.	Totals.
7	28	8	8	1	1	—	—	46
8	8	28	19	3	—	—	—	58
9	8	19	146	26	6	—	1	206
10	1	3	26	196	35	1	—	262
11	1	—	6	35	110	9	1	162
12	—	—	1	—	9	10	—	20
13	—	—	—	1	1	—	—	2
Totals	46	58	206	262	162	20	2	756

Now let us consider how to handle the material, allowing for the differentiation of the whorls. To begin with, our formula requires the use of the same number of homologous parts for each organism, and it is, on account of the value of the probable error of the random sample, undesirable to use fewer than 100 individuals. This leads to our cutting off Table V at the 10th whorl. In this way we get rid also of the forking, which certainly begins in many individuals at the 11th or 12th whorl. Table VIII gives us the data of Table V reconstituted for 110 plants, with ten whorls apiece. The only serious difficulty now remaining is that which I have referred to as arising from heterogenity in the first whorl. A glance at the mean and standard deviation of the branches in the first whorl given in Table V

Diagram 2.



Regression curve.—Whorl branches and position.

Table VIII.—Relation between Number of Branches and Position of Whorl.

Number of Branches in Whorl.

Position of Whorl.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	Totals.	Means.
1	2	2	3	8	8	6	9	24	20	17	11	—	—	110	7·7182
2	—	—	—	—	1	3	4	16	32	38	13	3	—	110	9·3273
3	—	—	—	—	—	—	7	6	29	42	21	5	—	110	9·7182
4	—	—	—	—	—	—	4	7	27	43	25	3	1	110	9·8273
5	—	—	—	—	—	—	6	7	28	39	28	2	—	110	9·7455
6	—	—	—	—	1	3	5	7	31	37	23	3	—	110	9·5636
7	—	—	—	3	2	5	9	19	27	28	16	1	—	110	8·8909
8	1	2	5	5	5	11	19	20	24	14	4	—	—	110	7·4909
9	5	10	11	12	8	14	19	16	10	5	—	—	—	110	5·7364
10	18	20	13	11	17	14	11	5	1	—	—	—	—	110	3·9636
Totals 2—10	24	32	29	31	34	50	84	103	209	246	130	17	1	990	8·2515

will suffice to demonstrate this heterogeneity. Certain individuals have the normal nnmber of about 8·5 branches to this whorl, but about a fifth of the total number of individuals only develop about half this normal number of branches. To illustrate this I have in Diagram 2 plotted the mean number of branches to the whorl, and fitted these means with a parabola of the third order,* using only whorls 2 to 10. The equation to this parabola is

$$y = 9{\cdot}451{,}443 - 0{\cdot}549{,}4302x - 0{\cdot}179{,}9881x^2 - 0{\cdot}008{,}4206x^3,$$

the origin being at the 6th whorl, and y giving the mean number of branches for x whorls from the 6th. We have the following results:—

Position of whorl.	Observed number of branches.	Calculated.
1	7·718	8·752
2	9·327	9·308
3	9·718	9·707
4	9·827	9·898
5	9·746	9·829
6	9·564	9·451
7	8·891	8·714
8	7·491	7·565
9	5·736	5·956
10	3·964	3·835

A much worse fit was obtained by striking a cubical parabola through all *ten* points.

It will be seen that the excellency of fit fully justifies the use of this curve. But that there is a large deviation from the observed mean of the 1st whorl, when we calculate its value from the curve thus obtained. Somewhat reluctantly, therefore, I felt compelled to omit the consideration of the 1st whorl from my investigations. Had I possessed a sufficient number of specimens I should have separated my material into two classes, those plants with normal 1st whorl and those with abnormal 1st whorl. But with my available material I should have had considerably less than 100 individuals to deal with, and accordingly I settled to take nine homologous parts only, namely, the 2nd to the 10th whorls, in which the differentiation appears to be solely due to position on the plant. Above the 10th whorl, the phenomenon of forking obscures the determination of branches to the whorl, while below the 2nd whorl the full or partial development of branches to the whorl seems to be determined by the local lower vegetation round the stem.

Taking Table VIII, I found for the mean of the means 8·2515 branches, and for the standard deviation of the means σ_M, $\sigma_M{}^2 =$

* By the method indicated in 'Biometrika,' vol. 2, p. 11.

3,938,354. Further, if σ be the standard deviation of the frequency distribution of branches, as found from the bottom row of Table VIII, we have

$$\sigma^2 = 6{\cdot}721{,}083.$$

Hence for use in formula (x) we have, since $m = 9$,

$$\frac{\sigma^2}{\sigma^2 - \sigma_M{}^2} = 2{\cdot}415{,}280, \quad \frac{1}{m-1}\frac{\sigma_M{}^2}{\sigma^2 - \sigma_M{}^2} = 0{\cdot}176{,}911 \text{ ...... (xi).}$$

Table IX gives the uncorrected homotyposis for the nine whorls treated as simple homotypes. From this I find,

$$\rho = 0{\cdot}131{,}258 \text{ .... (xii).}$$

Substituting (xi) and (xii) in (x), we find for the homotyposis of the number of branches in the whorls in *Equisetum arvense*, when corrected for differentiation due to position,

$$R = 0{\cdot}4939.$$

This result it must be admitted is extremely satisfactory, and indicates how it is quite possible to correct a result like (xii) by allowing for the differentiation of the homologous parts due to serial position.*

I hope before long to publish other results dealing with homotyposis in serial parts, where the differentiation has every variety of intensity. I think they will suffice to show that differentiation is not a subtle and evasive quality beyond the appreciation of the naturalist who is provided with the training requisite for modern biometric research.

(9.) The values of R as given by (ix) and (x) may be illustrated from the actual numbers for *Equisetum arvense*. We have seen in the footnote, p. 304, that

$$\eta = \sigma_M/\sigma.$$

This in our case gives

$$\eta = 0{\cdot}76549.$$

But by direct calculation on Table VIII, using whorls 2 to 10, Dr. Lee finds $r_{pc} = -0{\cdot}64616$. Hence with the notation of the footnote referred to

$$\sigma' = 0{\cdot}7632\,\sigma, \qquad \Sigma_M = 0{\cdot}4104\,\sigma.$$

* The value obtained for the crude homotyposis of the members of the whorls in *Asperula odorata* in my first memoir was $\rho = 0{\cdot}1733$ ('Phil. Trans.,' A, vol. 197, p. 326). I have little doubt that when we are able next summer to calculate the correction for differentiation in position of whorl, we shall find R for woodruff in good accordance with other homotypic results. My remarks about it were: "In counting the members on the whorls I soon found evidences of differentiation in position, the whorls towards the top of the spray having, as a rule, fewer members than those lower down" (*loc. cit.*, p. 325). Unfortunately I have not kept my records of position.

Table IX.—Uncorrected Homotyposis for *Equisetum arvense*.

Number of Branches in 1st Whorl of Pair.

Number of Branches in 2nd Whorl of Pair.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	Totals.
1	12	8	11	10	4	15	19	22	37	39	10	5	—	192
2	8	14	7	13	9	13	21	29	75	54	13	—	—	256
3	11	7	6	6	5	12	21	33	59	56	14	2	—	232
4	10	13	6	4	7	15	30	35	56	49	22	—	1	248
5	4	9	5	7	6	13	31	22	79	55	38	3	—	272
6	15	13	12	15	13	2	38	39	96	113	40	4	—	400
7	19	21	21	30	31	38	110	70	131	139	57	3	2	672
8	22	29	33	35	22	39	70	144	142	169	96	22	1	824
9	37	75	59	56	79	96	131	142	562	304	121	9	1	1672
10	39	54	56	49	55	113	139	169	304	770	197	21	2	1968
11	10	13	14	22	38	40	57	96	121	197	398	33	1	1040
12	5	—	2	—	3	4	3	22	9	21	33	34	—	136
13	—	—	—	1	—	—	2	1	1	2	1	—	—	8
Totals	192	256	232	248	272	400	672	824	1672	1968	1040	136	8	7920
Frequency of whorls	24	32	29	31	34	50	84	103	209	246	130	17	1	990

Thus η diverges much from r_{pc} and Σ_M from zero. Indeed, a glance at the diagram shows how far we are from true linear regression. If we use r_{pc} as above instead of η, *i.e.*, (ix) instead of (x), we have

$$R = 0{\cdot}3511,$$

a value very much below the actual value. This illustration will suffice to emphasise the importance of testing the actual curve of regression before we assume it to be linear and use equation (ix).

(10.) The subject of differentiation due either to position or age is, of course, a difficult one, but it does not seem at all beyond biometric treatment. The greatest difficulty which it seems to me will have to be encountered is not that of discovering and allowing for differentiation due to serial position, but in ensuring that when this has been allowed for, there is not remaining an organic correlation due to the necessity of adjacent parts "fitting." On this account it is most desirable that as large a number of homotypes as possible shall be taken, so that the part of the correlation due to the homologous parts having to fit, or, indeed, to serve a common end, should be reduced to as small a quantity as possible. For example, if we suppose adjacent whorls to have their number of branches influenced by an organic relationship, this result will only bias nine out of the forty-five pairs we should form in dealing with ten whorls. The question of separating organic from homotypic correlation is one that I hope to return to at a later date. Meanwhile the present paper will suffice to indicate how partial correlation coefficients enable the biometrician to free himself from the differentiation between individuals due to different periods of growth, or to different positions on the organism.

In conclusion I should like to thank Dr. Alice Lee and Mr. F. E. Lutz for aid readily granted at one or other stage of this investigation.