

ON THE LAWS OF INHERITANCE IN MAN*.

I. INHERITANCE OF PHYSICAL CHARACTERS.

BY KARL PEARSON, F.R.S., assisted by ALICE LEE, D.Sc.
University College, London.

CONTENTS.

	Page
(i) Introductory	357
(ii) Material. Family Record Series	358
(iii) Theory Applied	361
(iv) Size and Variability of Characters in two Generations	370
(v) Direct Assortative Mating in Man	372
(vi) Cross Assortative Mating in Man	375
(vii) Direct Parental Inheritance	377
(viii) Cross Parental Inheritance	383
(ix) Direct Fraternal Correlation	387
(x) Cross Fraternal Correlation	392
(xi) General Conclusions	393
(xii) Appendix of Correlation Tables	397

(i) *Introductory.*

ABOUT eight years ago I determined to supplement the data obtained by Mr Francis Galton for his work *Natural Inheritance* by a rather wider series of measurements on blood relations in man. Mr Galton had most generously placed his original data at my disposal and I had used them as far as stature was concerned in my memoir of 1895† and in a joint paper with Dr Lee in 1896‡. The eye-colour data of his Family Records were not reduced§ until after the discovery of a method for dealing with characters not capable of exact quantitative measurement||, and it is only recently that the full scheme of relationships back to great-grandparents has been completed¶. There were about 200 families in Mr Galton's records and only one measurable character, stature. The conditions

* I must gratefully acknowledge aid in the publication of the elaborate tables which accompany this memoir from a grant made to my department in the University of London by the Worshipful Company of Drapers.

† "Regression, Heredity and Panmixia." *Phil. Trans.* Vol. 187, pp. 253—318.

‡ "On Telegony in Man." *R. S. Proc.* Vol. 60, p. 274 *et. seq.*

§ "On the Inheritance of Eye Colour in Man." *Phil. Trans.* Vol. 195, pp. 102—121.

|| "On the Correlation of Characters not quantitatively Measurable." *Phil. Trans.* Vol. 195, pp. 1—47.

¶ F. E. Lutz: "Note on the Influence of Change in Sex on the Intensity of Heredity." *Biometrika*, Vol. II, pp. 237—240.

Biometrika II

46

as to age of the measured, or to method of measurement were not, perhaps, as stringent as might now be considered desirable, but Mr Galton's data were amply sufficient to lead him to his great discovery of the general form of the inheritance of blending characters in a stable community. The full significance of this discovery is hardly yet understood, and one constantly notices grave misinterpretations of Mr Galton's theory in the works of non-statistically trained biologists. The constants as determined from Mr Galton's stature data did not seem to me to be final; they were to some extent irregular and were not in full accord with the more uniform eye-colour results. It therefore appeared to me desirable to obtain further data, not only for several physical characters and to compare the results for these characters with those for mental characters, but to deal with both in as wide as possible a system of blood relationships. This was provided for in the following series of observations:

I. *Family Record Series.* About 1893 I drew up in conjunction with my then colleague, W. F. R. Weldon, the directions for family measurement which are described below. The measurements were in great part carried out by college students*, and I largely owe the success of this series to the energy and time devoted to the collection of the data by Dr Alice Lee. In the course of four to five years about 1100 cards were filled in. The tabling of the data on these cards and the calculation of the statistical constants, some 78 tables in all, are due entirely to Dr Lee, and occupied her spare time for nearly two years.

II. *School Record Series.* This series was started some years later and was aided by a grant from the Government Grant Committee. Its object was to record the mental and physical characters in pairs of brothers, of sisters, and of sisters and brothers in schools. About six thousand children were observed and measured, and provided more than 3000 pairs of brethren to illustrate in a great variety of ways the intensity of collateral resemblance in man. This series will only be dealt with incidentally in the first part of this paper, about 150 of the tables have been formed and the correlations deduced from them, but much work remains still to be done on the data for schools.

III. *Cousinship Series.* A third series on the ten kinds of first cousins is now being started with aid from the Government Grant Committee to complete our quantitative conceptions of collateral heredity. But it will be a number of years before the data here desired can be fully collected and still longer before the reductions can be completed. The above series form the material from which it is proposed to obtain quantitative measures of the degree of resemblance between blood relations in man. The present memoir deals primarily with the Family Record Series.

(ii) *Nature of the Family Record Series.*

It seems desirable to give the actual form of the instructions and schedule by aid of which the data were collected.

* I must take this opportunity of most heartily thanking the many helpers, who devoted much time and energy to measuring not only single but often 10 or 20 families.

FAMILY MEASUREMENTS.

Professor KARL PEARSON, of University College, London, would esteem it a great favour if any persons in a position to do so, would assist him by making one set (or if possible several sets) of anthropometric measurements on their own family, or on families with whom they are acquainted. The measurements are to be made use of for testing theories of heredity, no names, except that of the recorder, are required, but the Professor trusts to the *bona fides* of each recorder to send only correct results.

Each family should consist of a father, mother, and at least one son or daughter, not necessarily the eldest. The sons or daughters are to be at least 18 years of age, and measurements are to be made on not more than two sons and two daughters of the same family. If more than two sons or two daughters are easily accessible, then not the tallest but the eldest of those accessible should be selected.

To be of real service the whole series ought to contain 1000—2000 families, and therefore the Professor will be only too grateful if anyone will undertake several families for him.

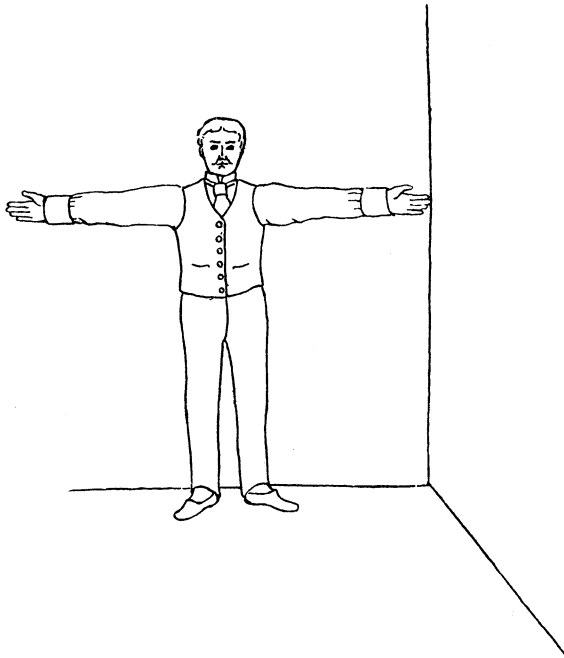
Copies of this paper, together with cards for recording data, may be obtained from

or from the above-named Professor.

The measurements required in the case of each individual are to be to the nearest quarter of an inch, and to consist of the following:—

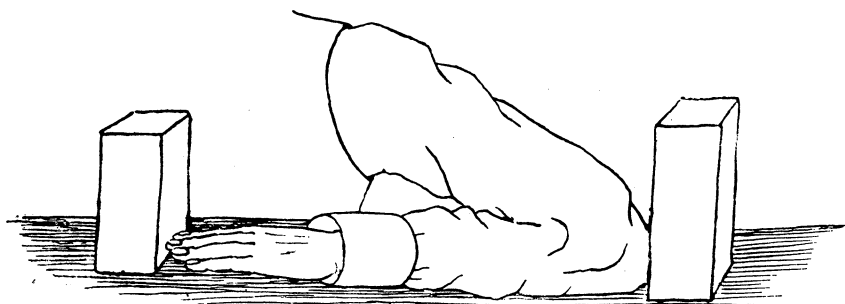
(I.) *Height*.—This measurement should be taken, if possible, with the person in stockings, if she or he is in boots it should be noted. The height is most easily measured by pressing a book with its pages in a *vertical plane* on the top of the head while the individual stands against a wall.

(II.) *Span of Arms*.—Greatest possible distance between the tip of one middle finger and the tip of the other middle finger, the individual standing upright against a wall with the feet well apart and the arms outstretched,—if possible with one finger against a doorpost or corner of the room.



On the Laws of Inheritance in Man

(III.) *The Length of LEFT Forearm.* The arm being bent *as much as possible* is laid upon a table, with the hand flattened and pressed firmly against the table, a box, book, or other hard object is placed on its edge so as to touch the bony projection of the elbow, another so as to touch the tip of the middle finger. Care must be taken that the books are both perpendicular to the edge of the table. The distance between the books is measured with a tape.



Or,

The arm being bent *as much as possible* the elbow is pressed against the corner of a room or the doorpost, the hand being flattened and pressed against the wall. The greatest distance from the tip of the middle finger to the corner or doorpost is to be measured.

Sample of filled in Data Card of Family Measurements.

		Height*		Span of Arms		Left Forearm	
		Feet	Inches	Feet	Inches	Feet	Inches
<i>One Family only</i>							
Father... .. (Not step-father)		5	9 $\frac{1}{4}$	6	1 $\frac{1}{2}$	1	7 $\frac{1}{4}$
Mother (Not step-mother)		5	0 $\frac{3}{4}$	5	2	1	4 $\frac{1}{4}$
	Age						
Son ...	26	5	7 $\frac{1}{4}$	5	11	1	6 $\frac{1}{2}$
Son ...	—	—	—	—	—	—	—
Daughter	30	5	4 $\frac{1}{4}$	5	5	1	4 $\frac{1}{2}$
Daughter	24	5	5 $\frac{1}{4}$	5	6 $\frac{1}{4}$	1	5

Name and Address of Recorder (not to be published in any way, but for convenience of reference).

**Miss A. L. Robinson,
Blounts Court Mansions, Kensington, S.W.**

.....
Both father and mother are absolutely necessary and should not be over 65 years of age.

All the measures are to be recorded to the nearest quarter of an inch. Before measuring read the notice circulated with this card, and kindly return the card as soon as possible to

[Name of individual collector was here inserted]
or to Professor Karl Pearson, University College, London, W.C.

* Put B against numbers if measure is taken in boots. If any person measured has ever broken a leg, arm or collar-bone, put L, A, C against all his or her measurements.

It is not for a moment suggested that the instructions or schedule form are ideal; they are of course open to criticism of a variety of kinds. But they were not settled without considerable thought and a definite reason for each point stated. Thus full growth is not reached at 18 years of age, perhaps not till 25. The growth, however, from 18 to 25 is relatively small, although sensible, and by fixing our limit at 25, we found a very large number of families would be cut off, for both parents would not be surviving, or, if surviving, beyond the age limit fixed for parents. Further, we should have been unable to interest college students effectively in the matter, as the bulk of them fall between 19 and 22. Again, it would have been better to take a lower maximum age for the parents, but in doing so we should again have greatly limited our available material. Better organs might undoubtedly have been selected than stature, span and forearm, e.g. head and finger measurements, but in such cases instruments and greater elaboration are needed, and the difficulty of obtaining upwards of 1000 families, already very great, would have been much intensified. We chose organs easily measured with moderate accuracy and asking for the nearest quarter-inch, only tabulated stature and span to the nearest inch, and forearm to the nearest half-inch. Thus the slight diurnal variations and the errors of measurement of the characters will not sensibly affect the constants calculated from our tables. Only a small percentage were measured in boots; we could not insist that ladies and gentlemen in middle life must remove their boots, or we might have met with a far larger number of refusals to be measured. Still the bulk of the measured did remove boots. After some experimenting on the effect of heels on apparent stature it was found that the subtraction of an inch from the recorded stature fairly represented the average increment due to boots. Hence the small percentage of boot entries was reduced before tabling by one inch.

Of course each family card did not provide us with four children, our maximum number allowed. Thus the number of our parental pairs lies for the different tables between 1000 and 1400, while for the fraternal correlations we have results based on 350 to 1400 pairs, according to the nature of the table. This is due to the fact that it was found far more difficult to get the measurements on *two* adult brothers, than on two sisters*. It was partly this defect in the number of pairs of brothers which led to the wider system of school measurements on brothers. The latter, however, do not modify but only confirm the results obtained from the smaller series in the *Family Records*.

I now propose to deal at length with the results obtained from our material.

(iii) *Theory applied.*

The regression in all cases is essentially linear, i.e. very closely linear within the limits of random sampling. It is impossible to give diagrams of all the 2×78

* Probably two adult brothers were far more rarely found both living at home, or if at home declined to be submitted to a measurement, which offered no immediate advantage to themselves.

The following is the Table for forearm in sons.

Forearm in Sons. 1156 Cases. Mean = 18".52, Standard Deviation = ".983.

Forearm in inches	14-14.5	14.5-15	15-15.5	15.5-16	16-16.5	16.5-17	17-17.5	17.5-18	18-18.5	18.5-19	19-19.5	19.5-20	20-20.5	20.5-21	21-21.5	21.5-22	22-22.5	22.5-23	23-23.5
Observed Frequency	1	1	1	5.5	10	39.5	95	177.5	260.5	225	166.5	105	41	14.5	7.5	4.5	—	—	1
Normal Frequency	.2		1.0	4.7	17.1	47.5	102.7	171.7	223.8	225.7	177.3	107.7	51.0	18.7	5.3	1.2	.2		

This Table as well as the previous one suggests that a small but sensible element of skewness in the forearm as well as the outlying group contributes to the divergence from normality.

It will be seen that our present data justifies Mr Galton's original use for stature of the normal curve and the normal surface, i.e.

$$z = \frac{N}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} e^{-\frac{1}{2} \frac{1}{1-r^2} \left\{ \frac{x^2}{\sigma_x^2} - \frac{2xyr}{\sigma_x\sigma_y} + \frac{y^2}{\sigma_y^2} \right\}}$$

(where $z \delta x \delta y$ is the frequency of a group of relative pairs having characters with deviations from their means lying between x, y and $x + \delta x, y + \delta y$; N being the total number of pairs, σ_x, σ_y being the standard deviations, and r the coefficient of correlation of the two characters: see *Phil. Trans.* Vol. 187, A, p. 264 *et seq.*). It also is fully justified for span and even for forearm (if we remember that there exists a small group of "outliers"). The normality of the distribution adds little, however, to our investigation, as long as we can show that the regression is practically linear (see Diagram III.). The practical value of normality arises chiefly when we pass from measurable characters in man to those that are not capable of exact quantitative measurement, for here every exception to normality weakens our general position.

The general linearity of our regression lines enables us in the present case to apply a simple theory, as soon as we have calculated the means, the standard deviations, and the correlations of the various characters.

These will enable us, by using the formulae of simple or multiple correlation, which depend simply on linearity, to predict the probable character in any individual from a knowledge of one or more parents or brethren ("siblings," = brothers or sisters). But without further assumption they do not enable us to test the effect of long-continued selection in establishing stocks; for we have no ancestral correlations, beyond the parental, for the characters dealt with. Ancestral correlations beyond the parental are, however, known for man in eye-colour inheritance (up to great-grandparents), for horses in coat-colour (up to great-great grandparents), and for dogs in coat-colour (up to grandparents). Hence, if the parental correlations for men, horses and dogs are sensibly the same,

we shall have small hesitation in assuming that the ancestral correlations for stature, span and forearm in man are closely alike in value to those for his eye-colour and for other characters in horse or dog. We shall thus be able to extend our theory, so as to deduce from our data the rate at which selection, natural or artificial, would establish stocks in man, and further, the limitations there are to the conception of an indefinitely active regression following on the suspension of selection.

It will be found that as far as the actual values are concerned our *Family Records* give values for heredity in man very sensibly larger than Mr Galton's stature data, and much closer to those obtained from his eye-colour data and for coat-colour in horses and dogs.

(iv) *Size and Variability of Characters in the two Generations.*

I will first consider whether there is a sensible change in type between the older and younger generation of our own epoch. The problem is not so easy to answer as it might *à priori* appear to some. We have the following results:

TABLE I. *Alteration in Type.*

MEANS		Stature	Span*	Forearm
1st Generation	{ Father ...	67''·68 ± ·06	68''·67 ± ·07	18''·31 ± ·02
	{ Mother ...	62''·48 ± ·05	61''·80 ± ·06	16''·51 ± ·02
2nd Generation	{ Son ...	68''·65 ± ·05	69''·94 ± ·06	18''·52 ± ·02
	{ Daughter...	63''·87 ± ·05	63''·40 ± ·05	16''·75 ± ·02
STANDARD DEVIATIONS		Stature	Span	Forearm
1st Generation	{ Father ...	2''·70 ± ·04	3''·14 ± ·05	0''·96 ± ·01
	{ Mother ...	2''·39 ± ·04	2''·81 ± ·04	0''·86 ± ·01
2nd Generation	{ Son ...	2''·71 ± ·04	3''·11 ± ·04	0''·98 ± ·01
	{ Daughter...	2''·61 ± ·03	2''·94 ± ·04	0''·91 ± ·01
COEFFICIENTS OF VARIATION		Stature	Span	Forearm
1st Generation	{ Father ...	3''·99 ± ·06	4''·64 ± ·07	5''·24 ± ·08
	{ Mother ...	3''·83 ± ·06	4''·62 ± ·07	5''·21 ± ·08
2nd Generation	{ Son ...	3''·95 ± ·06	4''·51 ± ·06	5''·29 ± ·07
	{ Daughter...	4''·09 ± ·05	4''·71 ± ·06	5''·43 ± ·07
ORGANIC CORRELATIONS		Stature and Span	Span and Forearm	Forearm and Stature
1st Generation	{ Father ...	·783 ± ·008	·752 ± ·009	·640 ± ·012
	{ Mother ...	·756 ± ·009	·677 ± ·011	·597 ± ·013
2nd Generation	{ Son ...	·802 ± ·007	·758 ± ·008	·686 ± ·011
	{ Daughter...	·828 ± ·006	·771 ± ·007	·716 ± ·009

* We note here a secondary sexual difference, the span on the average is about 1'' greater than the stature in man, and about ·5'' less than the stature in woman.

Now this Table contains a number of most interesting points.

In the first place the probable errors show us that for all three characters in both sexes the younger generation is distinctly larger than the older generation, son than father, daughter than mother. Is this a real progress in type? Taking Mr Powys' diagram for shrinkage in stature*, we should expect our men to reach a maximum at about 28 and our women at 25. Hence, since the average age of our younger generation is not more than 22 years, the younger generation cannot have reached its maximum. On the other hand, our average age of parents must be about 50. Let us suppose them to be 55 even. The difference in age of parents and offspring would thus mark a shrinkage of about '5" at a maximum. But the difference between fathers and sons is about an inch for stature and span, and for mothers and daughters about an inch and a half. It seems impossible therefore to attribute the whole change between the two generations to old-age shrinkage. In the next place, can it be due to periodic selection, i.e. only a portion of the younger generation become fathers and mothers? If so, we should expect not only a change in type, but a change in variability between the two generations. Comparing the standard deviations of fathers and sons, we see that fathers and sons are within the limits of random sampling equally variable. On the other hand daughters' standard deviations are in every case sensibly larger than those of their mothers. It would thus seem highly probable that the causes at work in the cases of the two sexes are not entirely the same. Mothers of adult children are a more stringently selected portion of the population than fathers appear to be. Of course some change in type between mothers and adult daughters is undoubtedly due to the fact of child-bearing, independent of any selection in childbed. But it is difficult to see how a physiological effect of this kind could change variability as well as type. I have shown that there is a slight correlation between size and fertility in women†, and this may be partially the source of the observed effect. Whether, however, the result be due to natural or reproductive selection, the change in the variability of the two generations of women seems to me to indicate that there is a selective change going on in the women of the middle classes in this country. The difference in type between fathers and sons,—since there is no change of variability,—might be more likely to be due to improved physical exercise. Of course a portion of the change in the women must also be attributed to this, but the change in variability forbids, I think, its being entirely attributed to this source.

However we judge the matter, whether we consider it due to selection, or to better environment, nourishment, or exercise, there seems no reason to suppose that the population, as far as the middle classes are concerned, is degenerating. In span, stature and forearm the younger generation is sensibly better than its parents.

* *Biometrika*, Vol. 1. p. 47.

† *R. S. Proc.* Vol. 59, p. 303. See also Vol. 66, p. 28 *et seq.*

If we compare the two sexes, we see that except in the matter of stature the married woman is *relatively* as variable as the married man, while in all three characters the young woman is relatively more variable than the young man. The supposed preponderance of male variability is thus again very fully negated, for large statistics of typical physical characters in mankind*.

Turning to the correlations we see (a) that in the older generation the mother is less highly correlated than the father, (b) that in the younger generation the son is less highly correlated than the daughter, (c) that the younger generation of both sexes is more highly correlated than the older generation. Now the effect of selection is to reduce correlation, hence if selection—a selective death-rate—be a real factor in the case of man and we know it to be so, we should certainly expect the correlations between the ages of youth and of middle life to be reduced. They are thus reduced, but far more markedly so in the case of woman than in that of man. Now as far as our data at present reach we know that the male baby is more variable and more highly correlated than the female†. In youth the woman is more variable and more highly correlated than the man; in adult age after child-bearing she is less highly correlated and perhaps very slightly less variable. It would thus seem that between birth and manhood the male is selected and falls in both variability and correlation below his sister. With womanhood comes her period of selection, sexual selection for wifeness, natural and reproductive selection for motherhood. These act with a little expected intensity and leave mothers of adult families with far less variability and correlation than their husbands have.

Of course these changes in variability and correlation may be partly growth changes, but since on the average the man reaches his maximum size four or five years later than the woman and at least four or five years beyond the average age of our group sons, it is difficult to account for the wide difference in variation and correlation between daughters and mothers as compared with that between sons and fathers by growth changes only.

I am inclined to think Table I. is very illustrative of the nature of selection among mankind, and further that it is also hopeful, not as regards the quantity, of which it takes no account, but as regards the quality of the offspring of a fair sample of the English middle classes.

(v) *Direct Assortative Mating in Man.*

We have seen above that all women, if they indeed become wives, do not become the mothers of adult children, i.e. the mothers of the second generation are not a random sample of their own generation. However it may arise there is

* See *The Chances of Death*, Vol. I. pp. 256—377. A recent criticism by Mr Havelock Ellis of my view that there is no preponderating variability of man over woman seems to need no reply, for the author does not appear to understand what weight is to be given to scientific evidence as compared with vague generalities.

† *R. S. Proc.* Vol. 66, p. 25.

certainly a "preferential mating"* taking place. I think we may safely assert that the first factor of sexual selection is active in man. I now turn to the second factor, "assortative mating." If certain women are rejected, at any rate as mothers of adult children, do the remainder mate at random as far as the above three characters are concerned? The answer is most decidedly in the negative, there is a very sensible resemblance in size between husband and wife, which *à priori* I should have said was hardly conceivable. Table II, gives the direct and cross coefficients for assortative mating in man.

TABLE II.

Assortative Mating. Based on 1000 to 1050 Cases of Husband and Wife.

	Husband's Character	Wife's Character	Correlation and Probable Error	Symbol
Direct	Stature	Stature	$\cdot 2804 \pm \cdot 0189$	r_{12}
	Span	Span	$\cdot 1989 \pm \cdot 0204$	r_{34}
	Forearm	Forearm	$\cdot 1977 \pm \cdot 0205$	r_{56}
Cross	Stature	Span	$\cdot 1820 \pm \cdot 0201$	r_{14}
	Stature	Forearm	$\cdot 1403 \pm \cdot 0204$	r_{16}
	Span	Stature	$\cdot 2023 \pm \cdot 0199$	r_{32}
	Span	Forearm	$\cdot 1533 \pm \cdot 0203$	r_{36}
	Forearm	Stature	$\cdot 1784 \pm \cdot 0201$	r_{52}
	Forearm	Span	$\cdot 1545 \pm \cdot 0203$	r_{54}

We see at once that between the same physical characters in the husband and wife of adult children there is a correlation of upwards of $\cdot 2$, a most remarkable degree of resemblance, greater than that of great-grandparents to their great-grandchildren (about $\cdot 19\ddagger$), and probably greater than that of first cousins to each other. We could hardly want stronger evidence of the existence of assortative mating in man, i.e. of the actuality of sexual selection. I had previously found \ddagger from Mr Galton's *Family Records*, that the correlation in stature between *husband and wife* was $\cdot 09 \pm \cdot 05$, but between *father and mother* of adult offspring was $\cdot 18 \pm \cdot 02$. Considering the comparative smallness of material, the latter result is in very good agreement with the present, but it seems to indicate that a portion of the observed resemblance in the parents of adult offspring is due to reproductive selection, i.e. homogamy being a factor of fertility. If the parents of adult children are on the average more alike than first cousins, then it follows that any evils which may flow from first cousin marriage depend not on likeness of characters, but on sameness of stock§.

That the whole result, further, is not due to a mere general custom of men and women mating with persons not differing widely from them in *stature*, is

* *Phil. Trans.* Vol. 187, p. 253 *et seq.* See especially p. 258.

† *Biometrika*, Vol. II, p. 221.

‡ *R. S. Proc.* Vol. 66, p. 30.

§ I have discussed this point more at length, *R. S. Proc.* Vol. 66, p. 29.

shown by the sensible correlation there is in eye-colour between husband and wife, i.e. $\cdot 10 \pm \cdot 04^*$, which is closely in agreement with the results for stature of husband and wife from the same data.

We may, however, estimate how far mating with regard to stature would produce resemblances in span and forearm. Let the subscripts 1, 3, 5 refer to three organs in a male of the population who marries, and 2, 4, 6 to the corresponding organs of a female of the marrying part of the population. Then $r_{13}, r_{35}, r_{51}, r_{24}, r_{46}, r_{62}$ are organic correlations such as we have tabled on p. 370. All correlations such as $r_{12}, r_{14}, r_{16}, r_{32}, r_{34}, r_{36}, r_{52}, r_{54}, r_{56}$ are zero, if we mated pairs at random. Now let them be assortatively mated and let $\rho_{12}, \rho_{34}, \rho_{56}$ represent the degree of resemblance in the sexual selection. Let r_{12}, r_{34}, r_{56} be the apparent correlations of mated pairs; then r_{12} will not be equal to ρ_{12} , for it is partly due to the degree of assortative mating indicated in ρ_{34} and ρ_{56} , because 3 and 5 are organically correlated with 1, and 4 and 6 with 2; thus the selection of 3's and 5's to associate with 4's and 6's would indirectly influence the relationship of 1 and 2, even if there were no direct associating of 1's and 2's. The relationship of r_{12}, r_{34}, r_{56} to $\rho_{12}, \rho_{34}, \rho_{56}$ may be easily found from my memoir on the influence of selection on variability and correlation†. We have only to put in the formulæ of pp. 15—17 the appropriate values for the population described above and we find:

$$\left. \begin{aligned} r_{12} &= \rho_{12} && + \rho_{34} r_{13} r_{24} + \rho_{56} r_{15} r_{26} \\ r_{34} &= \rho_{12} r_{13} r_{24} + \rho_{34} && + \rho_{56} r_{35} r_{46} \\ r_{56} &= \rho_{12} r_{15} r_{26} + \rho_{34} r_{35} r_{46} + \rho_{56} \end{aligned} \right\} \dots\dots\dots (i).$$

Now suppose that 1 and 2 represent statures, 3 and 4 spans, and 5 and 6 forearms. Then if all assortative mating be due to selection of stature, we might put ρ_{34} and ρ_{56} zero above and we should have:

$$\rho_{12} = r_{12}, \quad r_{34} = r_{12} r_{13} r_{24}, \quad r_{56} = r_{12} r_{15} r_{26}.$$

But $r_{13} = \cdot 7829, r_{24} = \cdot 7560, r_{15} = \cdot 6397, r_{26} = \cdot 5968,$ and $r_{12} = \cdot 2804$. This leads to

$$r_{34} = \cdot 1660 \text{ and } r_{56} = \cdot 1071,$$

as against the observed values:

$$r_{34} = \cdot 1989 \text{ and } r_{56} = \cdot 1977.$$

The former values are too small in both cases and, I think, we may safely assert, that the likeness of husband and wife in forearm and span is not *solely* due to a selection of stature.

Another explanation of these high coefficients of assortative mating has been suggested to me, namely that the population of England is built up of a number

* *Phil. Trans.* Vol. 195, A, p. 113. See also pp. 148—150, where it is shown that heterogamy rather than homogamy in eye-colour tends to increased fertility. If this be confirmed, eye-colour differs much in effect from stature.

† *Phil. Trans.* Vol. 200, p. 1 *et seq.*

of local races, and that men and women mate within their locality. Now it appears to me that this argument would be far more valid, if my material was drawn in bulk from local lower middle and artizan classes. But it is very doubtful how far it is true of the middle classes, such as provide the students at the London colleges. The middle classes undoubtedly marry in their own "sets," but these are hardly local sets. Further, a wide series of assortative mating observations have been made on another, wholly different class of characters, in which local race is regarded, and the coefficients come out as high as in the present data. Hence, I think, we are forced to the conclusion that the bulk of the observed resemblance in physical characters between parents is due to a direct, if quite unconscious, selection of like by like, and possibly in a contributory degree to a likeness in parents for the characters under consideration emphasising their fertility.

The amount of "consciousness" in the selection may possibly be measured by the difference between the stature-stature correlation and those for span-span, and forearm-forearm.

(vi) *Cross-Assortative Mating in Man.*

The second part of Table II. gives the cross-coefficients, for example, the correlation between husband's stature and wife's forearm. We might *à priori*, perhaps, anticipate that the correlation between a first organ in the husband and a second in the wife, would be equal to the correlation between the second in the husband and the first in the wife. This is actually the case for span and forearm, and, perhaps, we may consider for stature and span; the results for stature and forearm are less close than we might have anticipated, but the work has been revised without the discovery of any error. Relations such as:

$$r_{14} = r_{32}, \quad r_{16} = r_{52},$$

do not, however, appear to be theoretically necessary.

The problem now arises: are cross correlations between characters in husband and wife, solely due to selection of direct characters?

I think this may be roughly tested in the following manner. Suppose *only* these organs to be selected and the direct selection coefficients to be ρ_{12} , ρ_{34} , ρ_{56} , as before. They may be found from equations (i)* and we have the values:

$$\rho_{12} = \cdot 2374, \quad \rho_{34} = \cdot 0053, \quad \rho_{56} = \cdot 1043.$$

Thus there is most immediate selection of stature, a sensible selection of forearm, and practically none of span.

* These give numerically:

$$\cdot 2804 = \rho_{12} + \cdot 5919\rho_{34} + \cdot 3818\rho_{56},$$

$$\cdot 1989 = \cdot 5919\rho_{12} + \rho_{34} + \cdot 5087\rho_{56},$$

$$\cdot 1977 = \cdot 3818\rho_{12} + \cdot 5087\rho_{34} + \rho_{56}.$$

Now, if there be no *immediate* cross selection of other than these three organs and no immediate direct selection we should expect to find:

$$\left. \begin{aligned}
 \mathbf{r}_{32} &= \rho_{12}r_{13} + \rho_{34}r_{24} + \rho_{56}r_{35}r_{26} \\
 \mathbf{r}_{14} &= \rho_{12}r_{24} + \rho_{34}r_{13} + \rho_{56}r_{15}r_{46} \\
 \mathbf{r}_{52} &= \rho_{12}r_{15} + \rho_{34}r_{35}r_{24} + \rho_{56}r_{26} \\
 \mathbf{r}_{16} &= \rho_{12}r_{26} + \rho_{34}r_{31}r_{46} + \rho_{56}r_{15} \\
 \mathbf{r}_{36} &= \rho_{12}r_{13}r_{26} + \rho_{34}r_{46} + \rho_{56}r_{35} \\
 \mathbf{r}_{54} &= \rho_{12}r_{15}r_{24} + \rho_{34}r_{35} + \rho_{56}r_{46}
 \end{aligned} \right\} \dots\dots\dots (ii)$$

Substituting the ρ 's and the organic correlations in (ii) we find:

TABLE III.
Coefficients of Cross Assortative Mating.

Husband's Character	Wife's Character	Observed Value	Calculated Value
Span	Stature	·2023	·2327
Stature	Span	·1820	·2288
Forearm	Stature	·1784	·2171
Stature	Forearm	·1403	·2152
Span	Forearm	·1533	·1929
Forearm	Span	·1545	·1894

We conclude from this Table that: since the calculated values are all larger than the observed, the hypothesis that only direct selection of these three characters takes place is not valid. There must be direct selection of other correlated organs, or in some manner, as yet inexplicable, also an immediate cross assortative mating in man*. Generally, the results given in the present and the previous section for assortative mating, and in section (iv) for preferential mating, indicate that in future a greater degree of attention must be paid to sexual selection. It can hardly be so significant in the case of man, where most people would probably *à priori* suppose it of no account, and yet fail to play an important part in wild life. In particular, experimental enquiry on the relation of homogamy to fertility,—the likeness not being due to in-breeding—would be of very great value. It is clear, that “negative”† natural selection accompanied

* My own view, for which I have small evidence at present, is that the functions of sex are far more highly correlated with the physical characters in man than is generally supposed, and that the fertility of any given pair is very delicately atuned to the relative proportions of their frames. Hence when we come to correlate the physical characters in the parents of adult children, we find not only high direct but also high cross correlations, which seem inexplicable on any hypothesis of conscious assorting at mating. The only way to test this is to compare the correlations of husbands and wives at marriage with those of parents of adult children. We have already seen that these in certain cases sensibly differ. (See p. 373 and p. 374 footnote.)

† A selection for destruction not survival: see *Phil. Trans.* Vol. 200, A, p. 59.

by the correlation of homogamy and fertility would much aid us in comprehending the origin of species.

Although we are unable at present to account for the high coefficients of cross-assortative mating in man, it is possible to give an empirical formula, which will enable us to determine these coefficients in terms of the direct assortative mating coefficients and the organic correlations well within the limits of the probable errors of our results. Clearly the cross-assortative mating coefficients ought to vanish with both direct and organic correlations. Hence, if p, q refer to two organs in the husband and p', q' to the same pair in the wife, we should expect the cross correlation $r_{pq'}$ to be of the form:

$$r_{pq'} = Cr_{pp'}r_{p'q'} + C'r_{qq'}r_{pq}$$

Having satisfied myself that C and C' might be taken as practically equal, I found C as the mean of the last six entries in Table II. There resulted the formulae

$$\left. \begin{aligned} r_{pq'} &= \cdot5342 (r_{pp'}r_{p'q'} + r_{qq'}r_{pq}), \\ r_{p'q} &= \cdot5342 (r_{pp'}r_{pq} + r_{qq'}r_{p'q'}), \end{aligned} \right\} \dots\dots\dots (iii)$$

whence I found the following results.

TABLE III bis.

Calculated and Observed Cross Coefficients in Husband and Wife.

Husband's Character	Wife's Character	Observed Value	Calculated Value	Difference
Span	Stature	·202	·198	+ ·004
Stature	Span	·182	·196	- ·014
Forearm	Stature	·178	·159	+ ·019
Stature	Forearm	·140	·157	- ·017
Span	Forearm	·153	·151	+ ·002
Forearm	Span	·155	·151	+ ·004

The differences are well within the probable errors, and the above formulae may I think be safely used, if the cross coefficients are unknown.

(vii) *Direct Parental Inheritance.*

For the resemblance in like organs between offspring and parents we have for our three organs twelve cases. The correlations deduced from Appendix Tables XXII.—XXXIII. are given in Table IV. below.

It is impossible to regard these results without extreme satisfaction, not only as confirmation of the general reliability of the material, but also for the weighty evidence they bring for the nature of inheritance in man. When one remembers the labour of collecting the measurements, the days spent in tabling and reducing it, and the doubts which not unnaturally arose as to its value and the value of the tedious labour spent on it, the sense of satisfaction felt may be considered pardonable. The surprising agreement of the results—well within the probable

errors—for each character is to be noted in the first place. Considering that the measurements are made on more than 4000 individuals of different sexes in more than 1000 families, the conviction is complete that these numbers correspond to a

TABLE IV.

Coefficients of Heredity. Parents and Offspring.

Character	Father and		Mother and	
	Son	Daughter	Son	Daughter
Stature	$\cdot514 \pm \cdot015$	$\cdot510 \pm \cdot013$	$\cdot494 \pm \cdot016$	$\cdot507 \pm \cdot014$
Span	$\cdot454 \pm \cdot016$	$\cdot454 \pm \cdot014$	$\cdot457 \pm \cdot016$	$\cdot452 \pm \cdot015$
Forearm	$\cdot421 \pm \cdot017$	$\cdot422 \pm \cdot015$	$\cdot406 \pm \cdot017$	$\cdot421 \pm \cdot015$

reality in nature. From them we may safely draw the following conclusions for the organs examined :

(a) The son and daughter are equally influenced by their father, and equally influenced by their mother.

While a change of sex does appear to weaken hereditary influence in the eye-colour of man*, it does not appear to have any perceptible influence on the size of the human frame.

(b) In their influence on offspring there is no average prepotency of either father or mother, whatever there may be in individual cases.

(c) The inheritance of all characters does not appear to be the same.

The inheritance of forearm is for all four cases sensibly less than the inheritance of span, and that of span less than that of stature. We might as a probability put forward the following statement for further investigation.

(d) The more complex a character the greater the intensity of hereditary resemblance.

The fact that the correlation falls below $\cdot5$ with the simplicity of the character under consideration seems to suggest, however, that the reduction of the intensity cannot be due to an "alternative inheritance" in the case of the simple components of the character†.

For the mean values we have the following results :

Mean parental inheritance, father to son:	$\cdot463$
" " " " to daughter:	$\cdot462$
" " " mother to son:	$\cdot452$
" " " " to daughter:	$\cdot460$
Mean parental inheritance for both sexes and all characters:	460 .

* *Biometrika*, Vol. II. pp. 237—240.

† See *R. S. Proc.* Vol. 66, p. 142, and *Natural Inheritance*, p. 139.

I think we may fairly take the intensity of inheritance for measurable characters in man to be .46, or even for rougher work .5. It may be as well to put on record here the principal results for heredity in the direct line so far reached. I omit the results obtained in my memoir of 1895*, for I consider my present data to replace that series.

TABLE V.
Parental Inheritance in Different Species.

Species	Character	Mean Value	No. of Pairs used	Source	Remarks
Man	Stature	.506	4886	Present Memoir ditto ditto <i>Phil. Trans.</i> Vol. 195, p. 106	—
"	Span	.459	4873		—
"	Forearm	.418	4866		—
"	Eye Colour	.495	4000		—
Horse	Coat Colour	.522	4350	<i>Phil. Trans.</i> Vol. 195, p. 93	—
Basset Hound ...	Coat Colour	.524	823	<i>R. S. Proc.</i> Vol. 66, p. 154	Dams only used
Greyhound ...	Coat Colour	.507	9279	Unpublished data for two characters	Dams and Sires both used
Aphis (<i>Hyalopterus</i> } <i>Trirhodus</i>) ... }	Right Antenna Frontal Breadth	.439	368	<i>Biometrika</i> , Vol. I. p. 129) Ratios only taken to free from growth factor
Daphnia Magna ...	Protopodite Body Length	.466	96	<i>R. S. Proc.</i> Vol. 65, p. 154.	

I consider that this table contains the most reliable data we yet have collected and reduced for parental influence on offspring.

The general mean of the whole of these series is .48 and so far as we have yet gone, we may I think conclude, that:

(a) There is no reason for supposing parental heredity to be stronger in one species than a second.

(b) Its values lie between .42 and .52 and cluster round .48.

Thus for most practical purposes we may assume parental heredity for all species and all characters to be approximately represented by a correlation of .5.

In the course of the past 8 years many cases of parental inheritance have been worked out by the biometricians associated with me at University College, some of the most important of these are still unpublished, others have been replaced by far more reliable data; in further cases we know that the material was doubtful, e.g. the cephalic index for *fathers* and children of the North American Indians, or *sire* and offspring in the Basset Hounds. In such cases better material has been sought and our first results modified †. But in the present controversial phase of

* *Phil. Trans.* Vol. 187, p. 253 *et seq.*

† For example the greyhounds have shown that anomalies of the Basset Hound results were peculiar to the material, the cephalic index is inherited quite normally when we test it on material with reliable parentage, etc. etc.

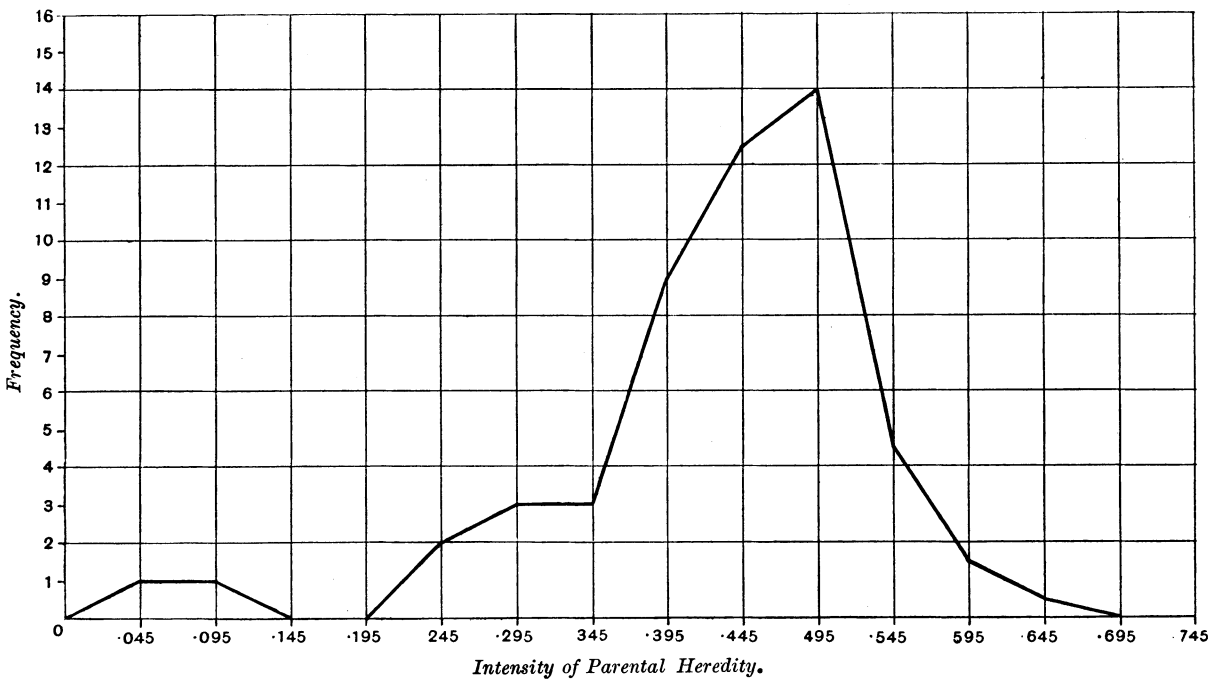
the theory of heredity, it seems well to hold no material back simply because one knows it to be untrustworthy. I therefore give the frequency distribution for every coefficient of parental correlation I am aware of, neither weighting them with the number of pairs on which they are based, nor remarking on the relative reliability of the data, which covers plants, insects and animals.

Frequency Distribution of Correlation Coefficients of Parental Heredity.

Magnitude of Correlation	.02—.07	.07—.12	.12—.17	.17—.22	.22—.27	.27—.32	.32—.37	.37—.42	.42—.47	.47—.52	.52—.57	.57—.62	.62—.67	Total
Frequency	1	1	—	—	2	3	3	9	12.5	14	4.5	1.5	.5	52

This distribution is represented graphically in the accompanying diagram; the mean and standard deviation of the system are $.430 \pm .010$ and $.107$ respectively. The fact of the cluster and its quantitative intensity are thus rendered obvious. The four extreme observations on the left are due to the Basset Hound Sires and North American Indian Fathers, both involving doubtful paternity*. If we omit

DIAGRAM VII. *Distribution of Correlation Coefficients in 52 Cases of Parental Heredity.*



* Both series are also very small, 100 to 400, as compared with the 1000 or more of most of the other series. As to their questionable character see *R. S. Proc.* Vol. 62, p. 414, and Vol. 66, p. 158 and especially footnote p. 159.

these as most certainly questionable the mean result is $\cdot453 \pm \cdot007$, the standard deviation being $\cdot071$. Thus $\cdot45$ may, I think, be justifiably taken in future to represent the approximate value of parental heredity, in cases where no direct observations have been made for the character and species under consideration. I prefer, however, the $\cdot46$ to $\cdot5$ of the *best* of the above series.

I now pass to the prediction formulae, i.e. the regression lines and planes, from which the probable value of a character in the offspring may be determined when the value of the character in the parentage is known.

If the subscript c denote child and p parent; and m be the mean, C the character; we have for prediction from one parent :

$$C_c = m_c + \frac{r_{cp}\sigma_c}{\sigma_p} (C_p - m_p) \dots\dots\dots(iv)$$

with a standard deviation for the array of value $\Sigma_c = \sigma_c \sqrt{1 - r_{cp}^2}$.

If we predict from two parents p_1 and p_2 , the formula is :

$$C_c = m_c + \frac{r_{cp_1} - r_{cp_2}r_{p_1p_2}}{1 - r_{p_1p_2}^2} \frac{\sigma_c}{\sigma_{p_1}} (C_{p_1} - m_{p_1}) + \frac{r_{cp_2} - r_{cp_1}r_{p_1p_2}}{1 - r_{p_1p_2}^2} \frac{\sigma_c}{\sigma_{p_2}} (C_{p_2} - m_{p_2}) \dots(v)$$

with a standard deviation for the array of

$$\Sigma_c = \sigma_c \sqrt{\frac{1 - r_{cp_1}^2 - r_{cp_2}^2 - r_{p_1p_2}^2 + 2r_{cp_1}r_{cp_2}r_{p_1p_2}}{1 - r_{p_1p_2}^2}}$$

Using these formulae we have the following results* :

A. *Stature.*

For Son :

- (1) Probable Stature = $33''\cdot73 + \cdot516$ (Father's Stature) $\pm 1''\cdot56$,
- (2) Probable Stature = $33''\cdot65 + \cdot560$ (Mother's Stature)† $\pm 1''\cdot59$,
- (3) Probable Stature = $14''\cdot08 + \cdot409$ (Father's Stature) $\pm 1''\cdot42$,
 $+ \cdot430$ (Mother's Stature) $\pm 1''\cdot42$.

For Daughter :

- (4) Probable Stature = $30''\cdot50 + \cdot493$ (Father's Stature) $\pm 1''\cdot51$,
- (5) Probable Stature = $29''\cdot28 + \cdot554$ (Mother's Stature) $\pm 1''\cdot52$,
- (6) Probable Stature = $10''\cdot82 + \cdot386$ (Father's Stature) $\pm 1''\cdot33$,
 $+ \cdot431$ (Mother's Stature) $\pm 1''\cdot33$.

* The actual tables of correlation are given in the Appendix and from them it will be seen that all possible pairs were used in each case for determining the correlation. Thus the standard deviations and means vary slightly from table to table, of course well within their probable errors. The formulae here given were, however, obtained by using the means and standard deviations which were adopted for Table I.

† If Father and Mother are to contribute indifferently to Son's stature, the parental statures should be in the ratio of about 560 to 516, which is very nearly the ratio of 1·085 to 1, and almost exactly equal to the 1·083 to 1 of ratio of Father's to Mother's average stature.

B. *Span.**For Son :*

- (7) Probable Span = $38''\cdot46 + \cdot458$ (Father's Span) $\pm 1''\cdot89$,
 (8) Probable Span = $38''\cdot38 + \cdot511$ (Mother's Span) $\pm 1''\cdot88$,
 (9) Probable Span = $18''\cdot04 + \cdot375$ (Father's Span)
 $+ \cdot423$ (Mother's Span) $\pm 1''\cdot70$.

For Daughter :

- (10) Probable Span = $34''\cdot20 + \cdot425$ (Father's Span) $\pm 1''\cdot77$,
 (11) Probable Span = $34''\cdot18 + \cdot473$ (Mother's Span) $\pm 1''\cdot77$,
 (12) Probable Span = $14''\cdot70 + \cdot355$ (Father's Span)
 $+ \cdot394$ (Mother's Span) $\pm 1''\cdot61$.

C. *Forearm.**For Son :*

- (13) Probable Forearm = $10''\cdot65 + \cdot430$ (Father's Forearm) $\pm ''\cdot60$,
 (14) Probable Forearm = $10''\cdot88 + \cdot463$ (Mother's Forearm) $\pm ''\cdot60$,
 (15) Probable Forearm = $5''\cdot58 + \cdot362$ (Father's Forearm)
 $+ \cdot383$ (Mother's Forearm) $\pm ''\cdot56$.

For Daughter :

- (16) Probable Forearm = $9''\cdot43 + \cdot400$ (Father's Forearm) $\pm ''\cdot56$,
 (17) Probable Forearm = $9''\cdot40 + \cdot445$ (Mother's Forearm) $\pm ''\cdot56$,
 (18) Probable Forearm = $4''\cdot50 + \cdot334$ (Father's Forearm)
 $+ \cdot371$ (Mother's Forearm) $\pm ''\cdot51$.

On the right is given in each case the probable error of the prediction*. We see from these formulae that with the selection of one parent only, the offspring rise to within 40 to 50 per cent. of the selected value; with the selection of *both* parents to within 70 to 80 per cent. of it. The diminution of the variability of the array due to two selected parents, is however only slightly less than that due to the selection of one parent only.

If we selected for two generations we should have offspring the same very nearly as the selected ancestry†. With our values for parental correlation, it is obvious that two or three generations of selection will suffice to bring the average of the offspring sensibly up to the selected ancestry, and the regression after this,

* In using these formulae for prediction, those not fully conversant with statistical theory, must bear in mind that they give only the mean or most probable results of a whole array of offspring due to *all* parents of definite characters. The validity of the formulae cannot be tested on merely individual cases. This warning is necessary because I have so often had *individual* cases in man or dogs cited as upsetting the whole of the ancestral law!

† "The Law of Ancestral Heredity," *Biometrika*, Vol. II, pp. 221—6.

if the stock mates with its like, will be very small or zero*. We cannot unfortunately on the present material determine absolutely its value; for, in the first place we have no correlations with grandparents or higher ascendants, and in the second place the assortative mating coefficients are so large, that we cannot afford to neglect them to a first approximation, as we have done for the case of eye-colour in man and coat-colour in horses†.

What, however, the present investigation impresses upon one is this: Parental correlation being from $\cdot45$ to $\cdot5$ in value is much higher than we could anticipate from Mr Galton's *Natural Inheritance* data. Hence selection is far more rapid in its effects than we supposed a few years ago, two to four generations suffice to effect what we originally considered would need 6 to 8. Further, the regression after such selection may well be zero. I have not worked out yet the multiple regression formulae allowing for assortative mating; they present considerable difficulty owing to the complexity introduced by the correlations between relations-in-law due to such mating. But neglecting for a moment the effect of assortative mating, the series of ancestral correlations,

$$\cdot46 \quad \cdot32 \quad \cdot23 \quad \cdot16$$

proceeding by a factor $\cdot7$ would give a zero-regression and not differ widely from the ancestral correlations we know for eye-colour in man‡. I lay no stress on these particular numbers, but I wish to emphasise the point that a few generations of selection in the case of man suffice to establish a breed, and that regression for this breed may well be insensible.

(viii) *Cross Parental Inheritance.*

I have defined *cross* heredity to be the correlation of two different organs in two blood relations§. We are now for the first time in a position to estimate its magnitude.

We see at once that these coefficients of cross heredity are for some cases almost as large as the coefficients of direct heredity, and on the whole sensibly larger than the values which but a few years ago were supposed to be those of the

* The physical aspect of this is perfectly easy to understand. When we select one parent the offspring advance 35 to 40 p.c. on the general population, when we select two parents 70 to 80 p.c. If we select two parents and four grandparents, there is another percentage increase which brings us up into the 90 per cents., and if we select for three generations we have nearly the 100 per cent. of the required character. Now suppose the selected stock to inbreed or otherwise mate with its likes for this special character. Why will the regression now that selection ceases be zero or insensible? For this simple reason, that while we cease to select within the stock, yet each new generation has an additional selected generation of ancestry behind it, and the influence of this ancestry balances the regression tendency. This is the simple verbal explanation of the cessation of regression with selection. The algebraical expression of its possibility was first given in my memoir of 1898: see *R. S. Proc.* Vol. 62, p. 401.

† *Ibid.* p. 224. See also *R. S. Proc.* Vol. 62, p. 388.

‡ *Biometrika*, Vol. II. p. 222. ϵ and ρ of p. 224 would be $\cdot56$ and $\cdot44$ respectively.

§ *Phil. Trans.* Vol. 187, A, p. 259. See also *R. S. Proc.* Vol. 62, p. 410.